

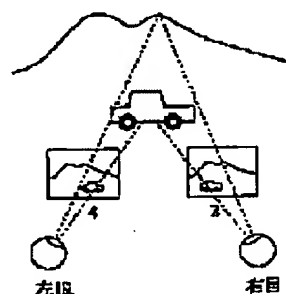
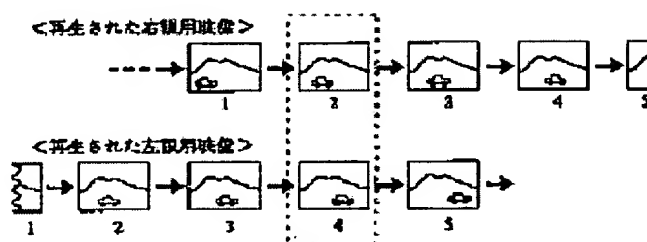
THREE-DIMENSIONAL DISPLAY IMAGE GENERATING METHOD

Patent number: JP8331607
Publication date: 1996-12-13
Inventor: MATSUMOTO YUKINORI; TERASAKI HAJIME;
 KATAYAMA MASAZUMI; SUGIMOTO KAZUhide;
 ARAKAWA TSUTOMU; SUZUKI OSAMU
Applicant: SANYO ELECTRIC CO LTD
Classification:
 - International: H04N13/04; G03B35/08; G06T7/00; G06T15/00
 - european:
Application number: JP19950257141 19950908
Priority number(s):

Abstract of JP8331607

PURPOSE: To provide the method for automation for generating an image for three-dimensional display with high precision by extracting the depth information of an original image to be the object of processing from that image and generating the image for three-dimensional display according to this information.

CONSTITUTION: The motion information of an object contained in the image is extracted. On this stage, the information is still handled as two-dimensional information. A display picture is taken on a coordinate plane and the motion of the object on this picture is described by a two-dimensional coordinate. Next, the relation of, correspondence between video frames is detected for grasping the motion of the object. Plural representative points are previously set to a video frame at time (t) [frame (t)] as a reference frame, the correspondent points of the respective representative points in the object frame at other time t' are pursued, and two-dimensional motion information is extracted. The relation of relative positions at every time is discriminated from the actual motion information of the object, and the depth information of each section is found from that relation. Parallax is decided corresponding to the depth information and images for both right and left eyes are generated. Generally, the closer the object is positioned, the more the parallax is enlarged. The right and left images are generated so that the parallax different for each section can not be generated.



Data supplied from the esp@cenet database - Patent Abstracts of Japan

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-331607

(43) 公開日 平成8年(1996)12月13日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 N 13/04			H 0 4 N 13/04	
G 0 3 B 35/08			G 0 3 B 35/08	
G 0 6 T 7/00			G 0 6 F 15/62	4 1 5
15/00		9365-5H	15/72	4 5 0 A

審査請求 未請求 請求項の数43 F D (全 24 頁)

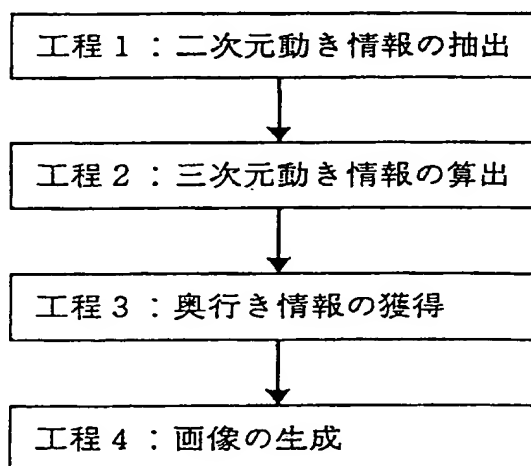
(21) 出願番号	特願平7-257141	(71) 出願人	000001889 三洋電機株式会社 大阪府守口市京阪本通2丁目5番5号
(22) 出願日	平成7年(1995)9月8日	(72) 発明者	松本 幸則 大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内
(31) 優先権主張番号	特願平7-96248	(72) 発明者	寺崎 肇 大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内
(32) 優先日	平7(1995)3月29日	(72) 発明者	片山 正純 大阪府守口市京阪本通2丁目5番5号 三洋電機株式会社内
(33) 優先権主張国	日本 (J P)	(74) 代理人	弁理士 吉田 研二 (外2名) 最終頁に続く

(54) 【発明の名称】 三次元表示画像生成方法

(57) 【要約】

【課題】 精度の高い三次元表示用画像を生成するための自動化に最適な方法が希求された。

【解決手段】 入力映像から、被写体の画面上の動き情報を抽出する(工程1)。このとき、ブロックマッチング等の手法を使う。つづいて三次元空間における被写体の現実の動きを算出する(工程2)。入力映像はもとの三次元的な動きの投影であるから、複数の代表点の動きから三次元表示の動きが逆変換によって求められる。この結果、被写体の座標が判明するため、被写体の奥行き情報が獲得できる(工程3)。この奥行きに従って視差を計算し、入力映像から両眼用画像を生成する(工程4)。



1

【特許請求の範囲】

【請求項 1】 処理の対象となる元の映像から該映像の奥行き情報を抽出し、この情報に従って三次元表示用の画像を生成することを特徴とする三次元表示画像生成方法。

【請求項 2】 請求項 1 に記載の三次元表示画像生成方法において、複数の映像フレーム間における映像各部位の二次元的な位置の変位を検出し、この情報から前記奥行き情報を抽出することを特徴とする三次元表示画像生成方法。

【請求項 3】 請求項 2 に記載の三次元表示画像生成方法において、前記複数の映像フレームは異なる時刻に撮影されたものであり、前記二次元的な位置の変位量に基づいてこれら複数の映像フレームを選択することを特徴とする三次元表示画像生成方法。

【請求項 4】 請求項 3 に記載の三次元表示画像生成方法において、前記映像各部位の二次元的な位置の変位量を統計的に処理し、その処理結果に基づいて前記複数の映像フレームを選択することを特徴とする三次元表示画像生成方法。

【請求項 5】 請求項 4 に記載の三次元表示画像生成方法において、前記変位量の統計的な処理は、前記映像各部位の動きベクトルの分散の導出であり、この分散が所定値よりも大きくなるように前記複数の映像フレームを選択することを特徴とする三次元表示画像生成方法。

【請求項 6】 請求項 5 に記載の三次元表示画像生成方法において、前記分散が所定値よりも大きくなるような複数の映像フレームを選択することができない場合には、三次元表示用画像の生成を中止することを特徴とする三次元表示画像生成方法。

【請求項 7】 請求項 2～6 のいずれかに記載の三次元表示画像生成方法において、前記二次元的な位置の変位から前記映像各部位が現実の三次元空間において占める相対位置関係を導出し、その結果に従って前記奥行きを決定することを特徴とする三次元表示画像生成方法。

【請求項 8】 請求項 2～6 のいずれかに記載の三次元表示画像生成方法において、前記二次元的な位置の変位から前記映像各部位の三次元的な動きを算出し、この動きから三角測量の原理によって前記映像各部位の三次元空間における位置座標を算出し、その結果に従って前記奥行きを決定することを特徴とする三次元表示画像生成方法。

【請求項 9】 請求項 2～8 のいずれかに記載の三次元表示画像生成方法において、

2

基準となる映像フレームに代表点を設定し、他の映像フレームにおける前記代表点の対応点を求め、これら代表点と対応点の位置関係を求めることによって、前記映像各部位の二次元的な位置の変位を認識することを特徴とする三次元表示画像生成方法。

【請求項 10】 請求項 9 に記載の三次元表示画像生成方法において、

前記基準となる映像フレームと前記他の映像フレームが異なる時刻に撮影されたものであるとき、

10 前記代表点と前記対応点の位置関係から、さらに他の映像フレームの対応点の位置を予想し、対応点を探索すべき領域を限定することを特徴とする三次元表示画像生成方法。

【請求項 11】 請求項 9、10 のいずれかに記載の三次元表示画像生成方法において、

前記代表点のうち、幾何的に特徴のある映像領域に関連するものについては、それら代表点の対応点の関連する映像領域についても前記幾何的な特徴が維持されるよう、前記対応点の位置を調整することを特徴とする三次元表示画像生成方法。

20 【請求項 12】 請求項 11 に記載の三次元表示画像生成方法において、

前記幾何的に特徴のある映像領域は、直線を含む領域であることを特徴とする三次元表示画像生成方法。

【請求項 13】 請求項 9 に記載の三次元表示画像生成方法において、

前記他の映像フレームにおける特定点近傍の映像領域と、前記基準となる映像フレームの代表点近傍の映像領域間で映像の類似性を評価し、この評価結果が良好な場合に前記特定点をもって前記代表点の対応点と決めることを特徴とする三次元表示画像生成方法。

30 【請求項 14】 請求項 9 に記載の三次元表示画像生成方法において、

前記他の映像フレームにおける特定点近傍の映像領域と、前記基準となる映像フレームの代表点近傍の映像領域間で映像の類似性を評価し、

前記特定点間の相対位置の妥当性を評価し、これら両評価の結果が良好な場合に前記特定点をもって前記代表点の対応点と決めることを特徴とする三次元表示画像生成方法。

40 【請求項 15】 請求項 14 に記載の三次元表示画像生成方法において、

前記各評価の結果を数値化して統合し、前記対応点の位置を変化させながら前記数値を再計算し、

繰り返し計算を通して前記対応点の位置精度を改善することを特徴とする三次元表示画像生成方法。

【請求項 16】 請求項 15 に記載の三次元表示画像生成方法において、

50 一旦全対応点の位置を固定した上で、1つの対応点のみ

を移動させながら前記各評価の結果が最良となる点を探索し、
探索された最良点の位置を前記 1 つの対応点の新たな位置とし、

これらの探索および位置の変更を順次全対応点について行うことを特徴とする三次元表示画像生成方法。

【請求項 17】 請求項 16 に記載の三次元表示画像生成方法において、

前記探索および位置の変更を全対応点について行った後、統合された各評価の数値が極値をとる条件を表すオイラー方程式を数値的に解くことにより、前記対応点の位置精度を改善することを特徴とする三次元表示画像生成方法。

【請求項 18】 請求項 13～17 のいずれかに記載の三次元表示画像生成方法において、

前記類似性の評価がブロックマッチングによって行われることを特徴とする三次元表示画像生成方法。

【請求項 19】 請求項 18 に記載の三次元表示画像生成方法において、

前記ブロックマッチングでは、比較すべき映像領域の間で色濃度の n 乗誤差総和 (n は 1 または 2) が計算されることを特徴とする三次元表示画像生成方法。

【請求項 20】 請求項 19 に記載の三次元表示画像生成方法において、

前記色濃度に対して、所定の色偏向定数を考慮して前記ブロックマッチングを行うことを特徴とする三次元表示画像生成方法。

【請求項 21】 請求項 20 に記載の三次元表示画像生成方法において、

前記色偏向定数は、前記 n 乗誤差総和が最小となるよう決められることを特徴とする三次元表示画像生成方法。

【請求項 22】 請求項 21 に記載の三次元表示画像生成方法において、

前記色偏向定数は、前記映像領域間の各画素の色濃度差の平均値であり、

前記ブロックマッチングは、前記映像領域間で各画素の色濃度差から前記色偏向定数を減算した後に 2 乗誤差総和を算出することを特徴とする三次元表示画像生成方法。

【請求項 23】 請求項 13～22 のいずれかに記載の三次元表示画像生成方法において、

映像に含まれる点を特徴点と非特徴点に分類し、特徴点である代表点について優先的に対応点を決めることを特徴とする三次元表示画像生成方法。

【請求項 24】 請求項 23 に記載の三次元表示画像生成方法において、

前記特徴点である対応点を補間して前記非特徴点である対応点を決めることを特徴とする三次元表示画像生成方法。

【請求項 25】 請求項 23 に記載の三次元表示画像生

成方法において、

前記基準となる映像フレームと前記他の映像フレームの間でブロックマッチングを行い、その結果、代表点との対応が良好である対応点を前記特徴点とすることを特徴とする三次元表示画像生成方法。

【請求項 26】 請求項 23 に記載の三次元表示画像生成方法において、

前記特徴点は、撮影時刻の異なる複数の映像フレームにおいて位置が安定的に変化している点であることを特徴とする三次元表示画像生成方法。

【請求項 27】 請求項 23 に記載の三次元表示画像生成方法において、

前記特徴点は、同時に撮影された映像フレーム間における位置の変位が、その近傍時刻において同時に撮影された映像フレーム間においてもほぼ一定であるか、または一定変化する点であることを特徴とする三次元表示画像生成方法。

【請求項 28】 請求項 1～27 のいずれかに記載の三次元表示画像生成方法において、

前記奥行き情報に従って両眼視差を決定し、この両眼視差に応じて左右両眼用の画像を生成することを特徴とする三次元表示画像生成方法。

【請求項 29】 請求項 2～27 のいずれかに記載の三次元表示画像生成方法において、

前記映像各部位の奥行き情報に従って部位ごとに両眼視差を決定し、この両眼視差に応じて部位ごとに左右両眼用の画像を生成することを特徴とする三次元表示画像生成方法。

【請求項 30】 請求項 29 に記載の三次元表示画像生成方法において、

前記映像各部位の奥行き情報に従って所定範囲の奥行きを持つ部位を選定し、選定された部位のみに対して左右両眼用の画像を生成することを特徴とする三次元表示画像生成方法。

【請求項 31】 請求項 28～30 のいずれかに記載の三次元表示画像生成方法において、

画像に対して前記両眼視差に応じた水平方向の変位を与えることにより、前記左右両眼用の画像を生成することを特徴とする三次元表示画像生成方法。

【請求項 32】 請求項 31 に記載の三次元表示画像生成方法において、

前記水平方向の変位を与えたとき、その影響によって画像枠に現れる段差を修正することを特徴とする三次元表示画像生成方法。

【請求項 33】 請求項 28～32 のいずれかに記載の三次元表示画像生成方法において、

前記奥行きが大きくなるに従って、前記両眼視差が次第に小さな値になるように決定することを特徴とする三次元表示画像生成方法。

【請求項 34】 請求項 28～33 のいずれかに記載の

5

三次元表示画像生成方法において、前記奥行きまたは前記両眼視差を所定の範囲に収まるように変換し、その結果得られた新たな奥行き情報に従って両眼視差を変更することを特徴とする三次元表示画像生成方法。

【請求項 3 5】 請求項 3 4 に記載の三次元表示画像生成方法において、

前記奥行きまたは前記両眼視差の広がり任意の点を中心として線形圧縮することにより、前記奥行きまたは前記両眼視差を前記所定の範囲に収まるように変換することを特徴とする三次元表示画像生成方法。

【請求項 3 6】 請求項 3 4 に記載の三次元表示画像生成方法において、

前記所定の範囲に収まらない奥行きまたは両眼視差を前記所定の範囲の上限値または下限値のうち近いほうの値に一律して変換することを特徴とする三次元表示画像生成方法。

【請求項 3 7】 請求項 3 4 に記載の三次元表示画像生成方法において、

変換後の値が前記所定の範囲の上限値または下限値に滑らかに収束する非線形変換を用い、前記奥行きまたは前記両眼視差が所定の範囲に収まるように変換することを特徴とする三次元表示画像生成方法。

【請求項 3 8】 請求項 2 8～3 2 のいずれかに記載の三次元表示画像生成方法において、

生成されるべき三次元表示画像における所定の領域の奥行き感を変更するために、その領域に与えるべき両眼視差を可変としたことを特徴とする三次元表示画像生成方法。

【請求項 3 9】 請求項 3 8 に記載の三次元表示画像生成方法において、

前記所定の領域の奥行き感を変更したとき、その変更に応じてその領域の表示面積を変更することを特徴とする三次元表示画像生成方法。

【請求項 4 0】 請求項 2 8～3 2 のいずれかに記載の三次元表示画像生成方法において、

映像の視点を仮想的に変更するために、想定される視点の変更に伴う映像各部位の位置の変位を前記奥行き情報から逆算し、前記位置の変位に従って映像を再構築することを特徴とする三次元表示画像生成方法。

【請求項 4 1】 請求項 4 0 に記載の三次元表示画像生成方法において、元の映像が 2 眼カメラによって撮影されたものであるとき、

これら 2 眼と所定の相関位置に仮想カメラを想定し、この仮想カメラから撮影されるべき映像を視点変更映像として生成し、

この視点変更映像と前記 2 眼カメラによって撮影された現実の映像とによって多視点映像を生成することを特徴とする三次元表示画像生成方法。

【請求項 4 2】 請求項 4 0、4 1 のいずれかに記載の

6

三次元表示画像生成方法において、

映像の視点を、ある映像フレームを撮影した視点から別の映像フレームを撮影した視点まで仮想的に移動させながら、移動経路上の任意の個所を視点とする視点変更映像を生成することを特徴とする三次元表示画像生成方法。

【請求項 4 3】 請求項 4 0～4 2 のいずれかに記載の三次元表示画像生成方法において、

前記再構築された映像について前記映像各部位の奥行き情報を再計算し、その結果得られた新たな奥行き情報から前記両眼視差を補正することを特徴とする三次元表示画像生成方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 この発明は三次元表示画像を生成する方法、特に、疑似的に三次元表示を行うために、単眼映像または多眼映像（ステレオ映像）から左右両眼用の画像を生成する方法に関する。

【0002】

【従来の技術】 単眼で撮影された二次元映像を元に疑似的に三次元表示を行う技術に、時間差方式の三次元表示がある。ここでまず、この原理を図 1～3 によって説明する。図 1 に示すように、背景は変化せず被写体のみが左から右へ移動する映像において、図 2 のように再生された右眼画像と左眼画像の間に一定の時間差を設けるとする。このとき、図 3 に示すような視差が生じる。ここで、図 2 と図 3 の数字はフィールド番号を示している。「視差」あるいは「両眼視差」とは、左右両眼からある点を見たときの視線ベクトルの角度差をいう。

【0003】 通常はこれら左右両眼の画像を、フィールドに同期して動作するシャッターメガネを介して見ることにより、立体視が可能となる。図 2 の場合、左眼の方がより進んだフィールドを見る。仮に被写体が逆向きに移動すれば、右眼に与えるフィールドを進ませればよい。こうした判断は例えば画像編集者によって行われる。以上が時間差方式の三次元表示の原理である。

【0004】 特開昭 60-263594 号公報には、この方式に基づく立体画像再生装置が開示されている。この装置は上記原理の適用において、

1. フィールドに同期した切換えを行うと、それぞれの眼に入る画像はインタレースの一方の画面のみとなって解像度が低下し、

2. 一方、切換えをフレームに同期させると、それぞれの眼に入るフリッカー周波数が 30 ヘルツとなってちらつきが増加する、という課題に鑑み、

1. 垂直・水平偏向信号を適倍して同一フィールドを複数回表示し、

2. その適倍周波数に合わせてシャッターメガネをより高速に切り換える、

という構成を採る。この構成によれば、両眼にインタレ

ースの両画面を表示しつつフリッカー周波数を高くすることができ、結果としてより自然な立体表示が可能となる。

【0005】さらにこの装置においては、被写体の動きに応じて時間差を1フィールドと2フィールドの間で選択する点について言及されている。動きが遅ければ大きめの時間差（つまり2フィールド）を持たせて立体視の効果を維持するというもので、かかる選択は「動き検出回路」によるものとされる。

【0006】なお、特開昭60-263594号公報に開示される技術を以降「従来技術」と呼ぶ。また本明細書においては、以後原則として、単に「映像」といえば処理の対象となる入力（ただし後述の視点変更映像を除く）、「画像」といえば処理後の出力（主に、最終的に左右両眼に表示される画像）を指す。

【0007】

【発明が解決しようとする課題】上記した従来技術によれば、時間差方式のもとで一定の立体視効果を得ることができる。しかし、以下の課題も存在する。

【0008】〔課題1〕技術的・機能的制約
従来技術では、時間差によって両眼視差を生じさせる必要性から、前提として、被写体が水平方向に動いていなければならない。これは根本的な制約である。さらに従来技術では、映像の組合せを出力するに留まるため、左右両眼用の画像（以下「左右画像」という）は既に撮影された映像から選択されなければならない。従って、

1. 動きのある映像を連続投入する必要がある、
2. 左右画像のために常時2枚の映像が必須となり、
3. 有限な組合せによって、三次元表示が不自然になることがあり、
4. その他、強調表示などの表現展開が困難である、等の問題が発生する。この結果、こうした三次元表示には自ずと限界があり、この限界は従来技術の応用と発展の足かせとなっている。

【0009】〔課題2〕処理自動化の困難性

従来技術によれば、「動き検出回路」によって処理の自動化が可能とされるものの、具体的な方途が不明である。またそれ以上に、もし課題1の解決を考えた場合、この動き検出回路程度の構成によって所望の画像を得ることはできない。すなわち、課題1を解決しつつ処理の自動化を図るとき、少なくとも以下の工程に関する具体的方法を示す必要がある。

【0010】1. 映像の解析

被写体の動き情報を獲得する。図1の例でいえば、被写体領域の移動方向と速度等、動き情報を自動的に抽出する工程

2. 画像の生成

前記解析結果に基づき、立体視に必要な視差を生む両眼用の画像を自動生成し、表示する工程

【本発明の目的】以上の諸点に鑑み、本発明の目的は以

下に約言される。

【0011】〔目的1〕技術的・機能的制約の解消
本発明では、既存の映像の組合せではなく、視点と被写体の距離情報（以降単に「奥行き情報」という）に従って新たに生成する。本発明は時間差という概念に拘束されない。

【0012】本発明で複数映像が必要なのは奥行き情報の抽出時のみであり、以降、1つの映像から自由に左右画像が生成される。これにより上記課題1の本質的解決を図るとともに、画像処理の自由度を向上させるものである。

【0013】〔目的2〕処理の自動化

目的1を達成しつつ、処理自動化のために以下の方法を具体的に開示する。

【0014】1. 映像の解析に当たり、被写体の任意方向の動きから奥行き情報を自動抽出する方法

2. 画像の生成に当たり、前記奥行き情報に従って被写体各部位に与えるべき視差を算出し、新たな画像を自動生成する方法

1. 2. の両手続の結果、時間差方式の前提条件である被写体の水平方向の動きは不要となる。

【0015】〔目的3〕多眼入力映像の三次元表示

目的1、2は主に単眼入力映像から三次元表示画像を生成する方法に関連する。そこで本発明は、目的1、2を達成する技術を利用して多眼入力映像を三次元表示するための画像生成方法も開示する。

【0016】

【課題を解決するための手段】本発明は、処理の対象となる元の映像から該映像の奥行き情報を抽出し、この情報に従って三次元表示用の画像を生成する。ここで「元の映像」とは、単眼で撮影されたものでもよいし、多眼で撮影されたステレオ映像でもよい。

【0017】このとき本発明は、複数の映像フレーム間における映像各部位の二次元的な位置の変位を検出し、この情報から前記奥行き情報を抽出する。「映像各部位」とは、映像を構成する各領域や各画素をいう。「映像フレーム」とは映像の処理単位で、通常の画像フレームの他、画像フィールド、MPEGのピクチャ等を含む。「複数の映像フレーム」は、単眼の場合は異なる時刻に撮影された複数フレーム（以下「異時刻フレーム」）をいう。多眼の場合は、同時に撮影された複数フレーム（以下「同時刻フレーム」）でもよいし、多眼カメラを構成する1つのカメラによる異時刻フレームでもよい。「二次元的な」とは、映像フレームが平面映像であることを意味する。多眼の場合でも、各カメラの映像は平面映像である。「二次元的な位置の変位」とは、平面上における位置の変位をいう。異時刻フレームなら時間経過に伴う位置の変化（すなわち動き）、同時刻フレームなら複数フレーム間の位置のずれを指す。

【0018】本発明のある態様では、前記複数の映像フ

フレームは異時刻フレームであり、前記二次元的な位置の変位量に基づいてこれらの異時刻フレームが選択される。変位量が大き過ぎる場合は、例えば時刻差を少なくして複数フレームを選択する方法が考えられる。

【0019】本発明では、前記映像各部位の二次元的な位置の変位量を統計的に処理し、その処理結果に基づいて複数の映像フレームを選択してもよい。統計的な処理として、映像各部位の動きベクトルの分散を導出し、この値が所定値よりも大きくなるように複数の映像フレームを選択してもよい。分散の導出は既知の計算によって行われる。映像各部位の動きが小さすぎると計算誤差が大きくなるため、分散がある程度大きいほうが好ましい。

【0020】従って本発明のある態様では、分散が所定値よりも大きくなるような複数の映像フレームを選択することができない場合、三次元表示用画像の生成を中止する。誤差の大きさに起因して不自然な三次元表示が行われる事態を回避するためである。

【0021】本発明は、こうして得られた映像各部位の二次元的な位置の変位から、映像各部位が現実の三次元空間において占める相対位置関係を導出する。その結果に従って前記奥行きを決定する。入力が単眼映像の場合、必ずしも絶対位置が求まらない。これは後述のスケールファクターの問題である。そこで、少なくとも相対位置を求める。

【0022】このとき、前記二次元的な位置の変位から前記映像各部位の三次元的な動きを算出し、この動きから三角測量の原理によって前記映像各部位の三次元空間における位置座標を算出し、その結果に従って前記奥行きを決定してもよい。単眼の場合はカメラの移動を仮定して三角測量の原理を適用する。多眼の場合は例えば左右映像に対してこの原理を適用する。

【0023】本発明では、複数の映像フレーム間における被写体の対応を把握する必要がある。そこで、基準となる映像フレーム（以下「基準フレーム」という）に代表点を設定し、他の映像フレーム（以下「対象フレーム」という）における前記代表点の対応点を求め、これら代表点と対応点の位置関係を求めて、前記映像各部位の二次元的な位置の変位を認識してもよい。これは一種の画像認識技術である。ここでいう「対応点」とは、代表点に対応する点である。対応点には、「真の対応点」と「計算上の対応点」の2つの概念が存在する。対応点は本来、各代表点に対して一意的に存在するはずであり、その存在個所以外の点が対応点であるという状態は考えられない。この理想的な対応点が真の対応点である。

【0024】一方、画像処理計算から決まる対応点は、真の対応点に一致するとは限らない。これが計算上の対応点である。計算上の対応点は、真の対応点の存在個所以外に存在することがありうるだけでなく、その位置

が適宜変更されることもある。後者は、例えば後述の対応点精度の改善処理を行う場合に発生する。

【0025】なお本明細書では、特に必要でない限り、「対応点」ということばを「真の対応点」と「計算上の対応点」の2つの概念に対して区別することなく用いる。

【0026】ここで、基準フレームと対象フレームが異時刻フレームであるとき、代表点と対応点の位置関係から、さらに別の対象フレームの対応点の位置を予想し、対応点を探索すべき領域を限定してもよい。最も簡単な例では、基準フレームと対象フレーム間で検出された被写体の動きが同じ速度で維持されると仮定して、別の対象フレームにおける被写体の存在個所を予想すればよい。

【0027】このとき、前記代表点のうち幾何的に特徴のある映像領域に関連するものについては、それら代表点の対応点の関連する映像領域についても前記幾何的な特徴が維持されるよう、前記対応点の位置を調整してもよい。「幾何的な特徴」を有する映像部分の例は、家の屋根の角や地平線などである。幾何的な特徴は被写体の幾何的特徴のみならず、映像中の色の変化点で構成されることもある。幾何的な特徴を維持すれば対応関係がより正確に把握されるという経験則に基づく。

【0028】このとき、幾何的に特徴のある映像領域として直線を含む領域を採用してもよい。基準フレームで直線をなした部分は対象フレームでも直線をなすと仮定するのである。対応点の位置を微調整して直線上に乗せる計算負荷は比較的軽い。

【0029】対象フレームにおける特定点近傍の映像領域と、基準フレームの代表点近傍の映像領域間で映像の類似性を評価し、この評価結果が良好な場合に前記特定点をもって前記代表点の対応点と決めてもよい。対応関係の把握に映像の類似性を採用するのである。「類似性」の判断基準は、色の近さ、形の近さなどである。

【0030】ここで「特定点」とは、評価に投入される対応点の候補をいう。候補の決め方は問わない。例えば等間隔に複数の特定点をとり、これらのうちで評価結果が最良だった点を対応点とすればよい。

【0031】さらに、この類似性だけでなく、特定点間の相対位置の妥当性を評価し、これら両評価の結果が良好な場合に前記特定点をもって前記代表点の対応点と決めてもよい。例えば、基準フレームで点Aよりも右にあった点Bが、類似性評価の結果、対象フレームにおいて点Aよりも左に存在するとされた場合、その周囲の対応関係にもこうした逆転現象があれば、点Bの位置は妥当と考えることができる。別の例としては、基準フレームに取られたある領域の重心は対象フレームでもその領域の重心付近になると仮定し、位置の評価を行ってもよい。

【0032】ここでは、2つの評価を終えてから対応点

11

を決めてもよいし、いずれか一方が終わった段階で対応点を仮りに決め、他方の評価によってその妥当性を判断してもよい。

【0033】こうした2つの評価を経て対応点が決まった場合、本発明のある態様では、各評価の結果を数値化して統合し、対応点の位置を変化させながらその数値を再計算し、繰り返し計算を通して対応点の位置精度を改善してもよい。つまり、2種類の評価結果を重み付けして加算し、この加算結果を最適化する対応点を求める。

【0034】このときさらに、一旦全対応点の位置を固定した上で、1つの対応点のみを移動させながら前記各評価の結果が最良となる点を探索し、探索された最良点の位置を前記1つの対応点の新たな位置とし、これらの探索および位置の変更を順次全対応点について行ってもよい。

【0035】ある対応点Xを移動するとき、他の対応点を固定した上で2つの評価を行い、その結果が最良となる点をもって対応点Xの新たな位置とする。これで対応点Xの1回の位置改善が終わる。つづいて、次の対応点について同様の探索と移動処理を行う。このときも他の点は固定しておく。処理は全対応点について行い、必要に応じて二巡以上繰り返す。以下、この改善手法を「固定探索手法」と呼ぶことにする。

【0036】固定探索処理が終了したのち、オイラー方程式を用いて対応点の位置精度を高めてもよい。オイラー方程式は2つの評価の統合後の数値が極値（極大値または極小値）をとる条件を表す。この数値には、全対応点の評価結果が反映されている。オイラー方程式によれば、画像の最小単位（例えば画素）よりも高い精度で対応点の位置を改善することも可能である。

【0037】本発明では、前記類似性評価をブロックマッチングによって行ってもよい。この方法によれば、色の類似性と形状の類似性の両方が同時に評価される。

【0038】より具体的には、このブロックマッチングは、比較すべき映像領域の間で色濃度の n 乗誤差総和を計算すればよい。 n は1または2とする。 $n=1$ のときは色濃度の差の絶対値の総和を求める。

【0039】ここで色濃度に対し、所定の色偏向定数を考慮してブロックマッチングを行ってもよい（以下「バイアスド・ブロックマッチング」と呼ぶ）。同時刻フレームの場合、複数のカメラの特性によって色濃度に一定の偏位が生じやすい。同じカメラによる異時刻フレームの場合にも、時刻によって天候が変わる（つまり映像の明度等が変わる）ようなとき、同様の現象が起こる。これを打ち消すように色偏向定数を与える。

【0040】ここで色偏向定数は、 n 乗誤差総和が最小となるよう決められる。そのように決めることで色濃度の偏位が最もよく解消されることが考えられるためである。

【0041】具体的には、色偏向定数は映像領域間の各画素の色濃度差の平均値とすればよい。このとき、最も

12

一般的な2乗誤差総和が最小になるためである。前記ブロックマッチングは、映像領域間で各画素の色濃度差から色偏向定数を減算した後に2乗誤差総和を算出すればよい。

【0042】本発明のある態様では、映像に含まれる点を特徴点と非特徴点に分類し、特徴点である代表点について優先的に対応点を決める。特徴点である対応点を補間して非特徴点である対応点を決めることもできる。基準フレームと対象フレームの間でブロックマッチングを行い、その結果、代表点との対応が良好な対応点を特徴点と決めてもよい。ブロックマッチングはバイアスド・ブロックマッチングでもよい。

【0043】本発明では、異時刻フレームにおいて位置が安定的に変化している点を特徴点としてもよい。別の言い方をすれば、動きベクトルが安定している点を特徴点とする。このような点は的確に追跡されていると考えられるためである。

【0044】同時刻フレーム間における位置の変位（すなわち位置のずれ）が、その近傍時刻において撮影された同時刻フレーム間においてもほぼ一定となるか、または一定変化する点を特徴点としてもよい。このような点も追跡の精度が高いためである。

【0045】こうして奥行き情報が獲得された後、本発明では、その奥行き情報に従って両眼視差を決定し、この両眼視差に応じて左右画像を生成すればよい。映像各部位ごとに両眼視差を決定し、左右画像を生成してもよい。左右画像は、両眼視差に応じた水平方向の変位を持つように生成すればよい。このとき、所定範囲の奥行きを持つ部位を選定し、選定された部位のみに対して左右画像を生成することもできる。例えば、奥行きが10メートル以内の映像部位のみを選定する場合、10メートルを超える領域（背景）に別の映像をはめ込むなどの加工が可能となる。

【0046】本発明のある態様では、水平方向の変位を与えたとき、その影響によって画像枠に現れる段差を修正する。例えば、画像端部に対して中央方向への変位が与えられたとき、その部分が中央方向に凹んで見えるためである。

【0047】両眼視差は例えば、奥行きが大きくなるに従って次第に小さな値になるように決定すればよい。両眼視差の定義に沿い、自然な三次元表示が可能となるためである。

【0048】このとき、奥行きまたは両眼視差を所定の範囲に収まるように変換し、その結果得られた新たな奥行き情報に従って両眼視差を変更してもよい。例えば、近すぎる奥行き（大きすぎる両眼視差）に修正を加えるためである。「所定の範囲」とは、例えば画像の奥行きが視点からディスプレイ画面までの距離よりも大きくなる範囲でよい。

【0049】変換の例として、奥行きまたは両眼視差の

広がりやを任意の点を中心として線形圧縮する方法、所定の範囲に収まらない奥行きまたは両眼視差をその範囲の上限値または下限値のうち近いほうの値に一律して変換する方法がある。後者の場合さらに、変換後の値が所定の範囲の上限値または下限値に滑らかに収束する非線形変換を用いてもよい。

【0050】本発明では、三次元表示用の画像が生成されたとき、画像の所定の領域の奥行き感を変更するために、その領域に与えるべき両眼視差を可変としてもよい。「所定の領域」とは、例えば人物の領域である。例えばこの人物の奥行きを大きくして（すなわち遠ざけて）見ることが可能となる。奥行きを小さくする場合は、一種の強調表示が可能となる。

【0051】ある領域の奥行き感を変更したとき、その変更に応じてその領域の表示面積を変更してもよい。例えば人物を遠ざけたいとき、この人物の奥行きを大きくすると同時に人物の面積を小さくすべきである。

【0052】本発明の別の態様では、映像の視点（すなわちカメラの位置）の仮想的な変更を行う（以下、こうして得られる映像を「視点変更映像」ともいう）。このために、想定される視点の変更に伴う映像各部位の位置の変位を前記奥行き情報から逆算し、前記位置の変位に従って映像を再構築する。これは、一種の画像加工技術である。例えば視点の高さを変えてみる場合、カメラの移動距離と前記奥行き情報から三角測量の原理等から被写体（映像各部位）の変位が計算できる。この計算結果に従って所期の映像を構築すればよい。このとき、再構築された映像について映像各部位の奥行き情報を再計算し、その結果得られた新たな奥行き情報から両眼視差を補正してもよい。

【0053】視点変更映像の応用技術として、元の映像が2眼カメラによって撮影されたものであるとき、これら2眼と所定の相関位置に仮想カメラを想定し、この仮想カメラから撮影されるべき映像を視点変更映像として生成し、この視点変更映像と前記2眼カメラによって撮影された現実の映像とによって多視点映像を生成してもよい。一般に多視点映像を得る場合、3眼以上のカメラによって映像を撮影しなければならない。そこでこの態様では、もともと2眼カメラによる映像が存在する場合、仮想的なカメラを別に設けたときに得られるべき映像を視点変更映像として生成し、これら3以上の映像によって多視点映像を生成する。例えば、この多視点映像のうち、2つの映像を選択して両眼視差を調整することにより、所望の立体表示を得ることができる。

【0054】別の態様として本発明は、映像の視点を、ある映像フレームを撮影した視点から別の映像フレームを撮影した視点まで仮想的に移動させながら、移動経路上の任意の個所を視点とする。例えば、映像フレーム1を撮影した視点を視点1、別の映像フレーム2を撮影した視点を視点2と定義すると、本発明は映像の視点を視

点1から視点2に仮想的に移動させながら、移動中の任意の位置を視点とする視点変更映像を生成する。映像フレーム1と2が異時刻フレームを構成する場合、視点変更映像を多めに生成すれば、映像フレーム1と2の間を滑らかに変化する映像を得ることができる。これは、例えば滑らかなスローモーション映像を得る場合などに適用することができる。スローモーションの各映像について必要な両眼視差の調整を行えば、この間も良好な三次元表示が実現する。当然ながら、映像フレーム1、2は同時刻フレームであってもよい。

【0055】

【発明の実施の形態】

実施形態1. 本発明の好適な実施形態を適宜図面を参照しながら説明する。実施形態1と2では、単眼で撮影された映像が入力されるものとする。多眼（ステレオ映像）の場合は実施形態3で説明する。

【0056】図4は本実施形態によって三次元表示画像を生成するための主な工程を示す図である。図に示すように、本実施形態は以下の工程に従って元となる二次元映像から三次元表示用の画像を生成する。ここで工程1～3が映像の解析、工程4が生成に関するものである。まず各工程の概略を説明する。

【0057】【工程1】二次元動き情報の抽出

映像に含まれる被写体の動き情報を抽出する。この段階ではまだ動き情報を二次元情報として扱う。表示画面を座標平面上にとり、被写体のこの画面上の動きを二次元座標で記述する。

【0058】本工程では、被写体の動きの把握のために映像フレーム間の対応関係を検出する。基準フレームである時刻 t の映像フレーム（以降「フレーム t 」という）に予め複数の代表点を設定しておき、別の時刻 t' の対象フレーム（以降「フレーム t' 」という）における前記各代表点の対応点を追跡する。フレーム t と t' は異時刻フレームを形成するが、これらは時間的に隣接するフレームである必要はない。また本工程の特徴は、被写体の水平方向のみならず、任意の方向の動きから二次元動き情報を抽出できる点にある。本明細書では、以降 t 、 t' がそれぞれ時刻を示す。

【0059】本明細書では以降「フレーム」といえば、フィールドなど広く映像構成単位全般を指すものとし、例えば525本の走査線によって構成されるテレビ受像器の1フレームや、640×480画素によって構成されるパーソナルコンピュータの一画面のみを指すものではない。また、代表点はフレーム t のみならず、フレーム t 、 t' の両方に設定してもよい。

【0060】【工程2】三次元動き情報の算出

被写体の二次元的な動きが判明すれば、この被写体の三次元空間における現実の動き情報（三次元動き情報）を算出する。この際、代表点と対応点の組を多数とることにより、被写体が実際に起こす動きを並進および回転の

両運動成分によって記述する。

【0061】 [工程3] 奥行き情報の獲得

被写体の現実の動きの様子がわかれば、各時刻における被写体の相対位置関係が判明する。この関係がわかれば被写体またはその各部位（以降単に「映像各部位」ともいう）の奥行き情報が求められる。

【0062】 [工程4] 画像の生成

奥行き情報に応じて視差を決定し、左右両眼の画像を生成する。この際、一般には近いものほど視差を大きくする。映像各部位で奥行きは異なるため、左右画像は部位ごとに異なる視差を生じるよう生成される。ここで注意すべきは、工程1における動き情報が任意方向の動きから抽出可能な事実と、本工程において視差をつける方向が原則として水平方向に限られる事実を混同してはならない点である。後者は画像を見る者の両眼が水平方向に位置することに由来する。なお、実施形態の説明においても「視差」または「両眼視差」は従来技術で行った定義に従うが、これ以外の定義であっても、容易に本発明の方法を適用することができる。

【0063】 以上が概略である。以降、各工程を詳述する。

【0064】 [工程1] 二次元動き情報の抽出

図5は映像フレーム間の対応関係を検出するためのフローチャートである。同図に示す各ステップを説明する。

【0065】 (S10) フレームtに代表点を設定

図6に示すように、まず基準フレームtに代表点を設定する。同図では、フレームtを8×8画素ごとにメッシュによって分割し、各代表点を格子点に置いている。ここでは左からi番目、上からj番目の代表点をPt(i, j)と記述し、Pt(i, j)に対する時刻t'の対応点をPt'(i, j)と記述する。また、必要に応じてPt(i, j)のx, y座標をそれぞれ、Pt(i, j)x, Pt(i, j)yと記述する。

【0066】 なお本ステップにおいて、代表点は格子点に限らず任意の箇所に配置してもよい。極端な場合は全画素を代表点とすることもできる。

【0067】 (S11) 対応点候補領域の設定

例えば図6に示すPt(6, 4)という代表点を考えた場合、Pt'(6, 4)の存在しうる領域を予め設定する。これは映像の動きがある限度以上に急激な場合を除き、Pt'(6, 4)がPt(6, 4)の近傍にあるという仮定に基づく。本実施形態では、例えばPt(6, 4)の近傍100×60画素の領域にPt'(6, 4)が入るものとし、Pt'(6, 4)の検出の際の計算量を削減する。

【0068】 本ステップについては以下の応用が可能で*

$$E1 = \sum \sum |It(Pt(i, j)x+u, Pt(i, j)y+v) - It'(Pt'(i, j)x+u, Pt'(i, j)y+v)|^2 \quad (式1)$$

によって求められる。ここで2つのΣはu, vに関する。これらはそれぞれ、

*ある。

【0069】 1. 映像が比較的激しく動いているときはフレームt, t'が隣接するようにt'を決める。代表点の位置変化は最小に抑制され、対応点が前記領域に入らないおそれも最小となる。ただし当然ながら、予め候補領域を画面全体とする方法も考えられる。その場合、計算量が増加するが、映像の大きな動きによって対応点を取り逃すおそれは減る。

【0070】 2. 本実施形態では単にPt'(6, 4)がPt(6, 4)の近傍にあると仮定したが、いったんPt(6, 4)の複数フレームにおける移動軌跡が判明したときは、この軌跡の延長線上に候補領域を決めることも可能である。映像の動きがある程度一定している場合、この方法による対応点候補領域の絞り込みは非常に有効である。

【0071】 (S12) 対応点候補領域における非類似度の計算

つづいて、候補領域から対応点の位置を具体的に決めていく。ただし本ステップでは、前ステップとは逆に、映像の動きが緩慢過ぎる場合に問題が生じる。動きが少ないと動き情報の抽出が困難になり、抽出された情報が大きな誤差を含む可能性があるためである。

【0072】 そこでこのようなときには、予めフレームt, t'がある程度離れるようt'を決める。この際、映像各部位の変化量を統計的に処理し、例えば変化量の分散が所定値を超えるようなt'を決めればよい。なお、仮に条件を満たすt'が見つからなければ、三次元表示用画像の生成を中止する。このときは入力された映像をそのまま出力するか、または映像の全部位に同じ奥行きを付与して表示すればよい。

【0073】 本ステップでは、対応点の位置を決めるべく、フレームt, t'間でブロックマッチングによる非類似度の計算を行う。前記対応点候補領域内のある点を中心とする近傍ブロックと代表点の近傍ブロックとの間で濃度の2乗誤差の総和（非類似度）を求め、これが最小になる点を計算上の対応点と決める。

【0074】 図7はブロックマッチングの様子を示す図で、本実施形態では9画素で1ブロックと定義し、中央の画素を代表点とする。

【0075】 同図において、まずフレームt上にPt(i, j)を含むブロック1を取り、フレームt'側では対応点の仮の候補Pt'(i, j)を含むブロック2を取る。ここで一般に、画素(x, y)の時刻tにおける画素値をIt(x, y)と記述すれば、非類似度(E1とする)は、

【数1】

$$u = -1, 0, 1$$

$$v = -1, 0, 1$$

17

の値を取り、仮の $Pt'(i, j)$ に対して合計 9 画素に関する 2 乗誤差総和が計算できる。そこで $Pt'(i, j)$ を候補領域内で少しずつ動かしていき、 $E1$ が最小になる点を対応点と決めればよい。

【0076】図 8 は $Pt'(i, j)$ ごとに $E1$ の値を縦方向に表した模式図であり、同図の場合、非類似度が急峻なピークをとる点 Q が対応点と決まる。以降、同様に他の代表点についても対応点を決めていく。

【0077】なお、本ステップには以下の応用または変形がある。

【0078】1. ここでは濃淡画像を前提に、濃度の 2 乗誤差を計算したが、カラー画像の場合、RGB 各濃度の 2 乗誤差の総和、すなわち $E1_R + E1_G + E1_B$ を非類似度としてもよい。これは他の色空間による濃度、例えば HVC 濃度であってもよい。2 乗誤差の代わりに単なる誤差の絶対値（残差の総和）を採用してもよい。
2. 本ステップでは 1 ブロックを 9 画素としたが、通常はある程度多数の画素によってブロックを定義することが望ましい。例えば通常のパーソナルコンピュータまたはワークステーションの高解像度画面を想定する場合、実験の結果、 16×16 画素程度のブロックによって良好な結果が得られている。

【0079】(S13) 対応点の初期位置の決定
前ステップにより、一応は対応点を求めることができた。しかしこの段階では対応点の位置が正しいとは限らない。被写体の境界やエッジについては対応点が比較的正しく求まるものの、例えば変化の乏しい映像部位では、対応点の位置は相当誤差を含んでいるとみるべきである。図 8 でいえば $E1$ が明確なピークをとらない状態といってよい。図 9 は前ステップの結果求められた対応点と代表点の関係を示す図で、同図に示すごとく、家や木、特にそれらの輪郭等の特徴的な点については対応点的良好に求まるが、空や地面については誤差が大きい。

【0080】そこで本ステップと次ステップにより、対応点の位置補正を行う。本ステップでは、まず対応点の初期位置という概念を導入し、実際にその初期位置を決める。つづいて、次ステップで繰り返し計算による位置精度の改善を図る。

【0081】本ステップにおける対応点初期位置の決定には以下の方針が考えられる。

【0082】1. 前ステップで求められた全対応点を同等に扱う

全対応点の位置をそのままそれらの初期位置として次ステップへ投入する。

【0083】2. 対応点の扱いに差異を設ける

当初よりある程度正しい位置にあると思われる対応点

(以降「特徴点」という)の位置はそのまま初期位置と*

$Pt'(i-1, j)$ 、 $Pt'(i+1, j)$ 、 $Pt'(i, j-1)$ 、 $Pt'(i, j+1)$

が隣接している。 $Pt'(i, j)$ は通常、これら 4 つの点の重心付近に存在すると考えるのが妥当である。これ

18

*として使い、そうでない対応点(以降「非特徴点」という)の初期位置については、前記特徴点のそれをもとに決めるというものである。ここで特徴点としては以下の点を想定することができるが、これらの点は現実には一致すること多い。

【0084】(1) 前ステップの $E1$ が明確なピークを示した対応点

このような対応点の位置精度は一般に高いためである。

【0085】(2) 直交エッジ成分が多く存在する箇所にある対応点

建物の角などの部分では対応点の位置がかなり正しいと思われるためである。

(3) フレーム t 、 t' 、...において位置が安定的に変化している対応点

ここで変化の安定性は、動きベクトルの一定性ということができる。ここでは、フレームの進行に伴って移動方向、移動距離が一定な対応点を選ぶ。具体的には、例えば動きベクトルのばらつきが所定値以下の対応点を選定する。こうした対応点は的確に追跡されているはずであり、代表点と正しい対応関係にあると判断できるためである。ただし、例えば映像を撮影するカメラが不規則に動いた場合には、その影響を加味して判断する。

【0086】こうして特徴点を選定されれば、これらはそのまま初期位置として使い、非特徴点の初期位置は特徴点の位置を補間したり、または特徴点の近傍から順に決めていくことができる。すなわち、前ステップによる非特徴点の位置精度は低いため、精度の高い特徴点から幾何的に非特徴点の初期位置を与えるのである。なお当然ながら、(3)の特徴点を見つける場合にも前ステップの方法を有効利用できる。

【0087】以上、特徴点の選定に基づいた対応点初期位置の決定方法を述べたが、この他、動的計画法(ダイナミック・プログラミング)を用いて対応点の初期値を求めてもよい。

【0088】(S14) 対応点の改善処理

対応点の位置の妥当性を評価するために式を導入し、繰り返し計算によって位置精度を改善する。S12 ステップでは非類似度を評価する式 1 を導入したが、ここではさらに、対応点間の相対位置関係の妥当性を評価する式を導入し、これら 2 つ評価結果を統合して位置の改善を図る。

【0089】図 10 は相対位置を評価する原理を説明する図である。同図において、各点はそれぞれ対応点を表している。このうち、図中の $Pt'(i, j)$ を中心に考えると、これには以下の 4 つの対応点、

【数 2】

は、映像各部位が動いても画素単位の微視的な見方をすれば相対位置関係がほぼ保たれるという経験則に基づい

10

20

30

40

50

19

ている。なお、この性質を数学的にいえば、 (i, j) の関数 $Pt'(i, j)$ の 2 次微分がほぼ 0 であるということにほかならない。

$$E2 = |Pt'(i, j)x - St'(i, j)x|^2 +$$

が相対位置の妥当性評価式となる。この式だけを考えれば、 $E2$ が最小値になるときに対応点の位置が最も望ましい状態となる。

【0091】本ステップでは、式 1 および式 2 の評価結果を適当な結合定数 k で加算し、

$$E = E1 / N + k \cdot E2 \quad (\text{式 3})$$

で表される E を最終的な評価式とする (N はブロックマッチングの際に定義された 1 つのブロックに含まれる画素数である)。すなわち、まず各対応点について E を計算し、続いて全対応点の E の総和 ΣE を計算し、 ΣE が最小値となるよう、各対応点の位置をすこしずつ変化させる。 ΣE の値が収束するか、または繰り返し計算を所定の上限回数に達するまで行い、改善処理を施す。より具体的には、各対応点の位置を変化させるとき、以下のいずれかの方法を実施すればよい。

【0092】(1) オイラー方程式を解く方法

ΣE が極値 (ここでは極小値) をとる条件を示すオイラー方程式を数值的に解くことによって対応点を得る。この手法自体は既知である。これは、各代表点を含むブロックでの画像傾き情報と、対応ブロック間の画素差分情報から改善すべき方向を見出し、これに基づいて対応点の位置を初期位置から徐々に動かしていき、最終解を求める。

【0093】(2) 固定探索手法

まず、対応点候補領域において、改善対象の対応点の E が最小になる点を探し、これを新たな対応点とする。このとき、他の点の位置を不動とみなして探索を行う点に特徴がある。この処理を順次全対応点に対して行う。

【0094】(3) 混合手法

(2) の手法によれば、画素単位の精度で対応点の位置が求まる。一方、(1) によれば、理論上画素単位以下の精度で位置を求めることができる。そこで、まず

(2) の手法によって画素単位の精度で対応関係を求め、しかる後にオイラー方程式を適用して精度を高めることも可能である。

【0095】なお実験によれば、同じ精度で比較した場合、(2) の手法によって (1) よりも短時間で好ましい解が得られている。

【0096】図 11 は図 9 の対応点候補に対して本ステップの改善処理を行った結果を示す図である。実験によれば、カラー画像の場合、 k は 5 ~ 200 程度で良好な結果が得られることがわかった。図 9 と図 11 はともに模式的な図であるが、実験の結果、実際にこれらの図に近い改善が見られた。

【0097】以上が工程 1 の詳細である。本工程の特徴は被写体の任意方向の動きから二次元動き情報を抽出で

20

* 【0090】従って上記 4 点の重心を $(St'(i, j)x, St'(i, j)y)$ と表記すれば、

【数 3】

$$|Pt'(i, j)y - St'(i, j)y|^2 \quad (\text{式 2})$$

きる点にある。これは代表点と対応点という概念で動きを把握する利点であり、水平方向の動きを検出して時間差を決定する従来技術に比べ、広い応用を可能とするものである。

【0098】なお本工程には、以下の応用または変形がある。

【0099】1. $E2$ 導出の際、上下左右の 4 点のみならず、斜め方向の 4 点を加えた計 8 点の重心を考える。いかなる組合せが最適であるかは映像の種類にも依存するため、適宜実験によって決めていくことが望ましい。

【0100】2. 式 3 による評価は、 $E1$ のみによる評価結果が思わしくなかった対応点から優先的に行う。これは $E1$ の結果が悪い対応点は一般に位置の誤差が大きいと考えられるためであり、こうした対応点の位置を早期に、かつ大幅に改善することが望ましいためである。

【0101】3. 位置改善の際、幾何情報も利用する。フレーム t において幾何的に特徴のある領域、例えば直線を形成していた複数の代表点については、それらの対応点も直線を形成するように位置を補正する。これは映像上直線に見える部分は現実の三次元空間でも直線である可能性が高く、一方、三次元空間の直線はフレーム t' でも直線となるべきだからである。本来奥行きは直線に沿って一様に変化するものであり、直線に沿う変化は視覚的に容易に把握されるため、この方法による改善効果は大きい。かかる改善をしない場合、最終的に生成される画像において直線に沿う奥行きにむらが生じ、不自然な三次元表示になることがある。

【0102】4. さらに別のフレームについても対応点を求める。本工程ではフレーム t に対するフレーム t' の対応点を求めたが、さらに第三のフレーム t'' における対応点も求め、映像各部位の平均化された動きを求めることができる。この方法は、フレーム t' における対応点位置を改善していくのではない。多くのフレームで対応点をとることにより、対応点の位置とそのフレームが撮影された時間から映像各部位の動きを統計的に決めていくものである。

【0103】【工程 2】三次元動き情報の算出

工程 1 により、映像各部位の画面上の二次元的な動きが判明した。工程 2 ではこの情報から各部位の三次元的な動きを算出する。映像は被写体の現実の動きを平面に投影したものであり、本工程では代表点と対応点の位置関係からもとの動きを導出する。

【0104】一般に被写体の三次元空間における動きは、並進運動と回転運動の合成として記述することができる。ここではまず、動きが並進運動のみで構成される場合の計算方法を説明し、後に一般化された方法を概説

する。

【0105】1. 動きが並進運動のみの場合

図12はある点Pの画面上の移動と三次元空間での現実の移動の対応を示す図である。同図では画面上の二次元座標を大文字X等で、現実の三次元座標を小文字x等で表記するものとし、三次元座標のうちx、y軸を画面上に、z軸を奥行き方向にとっている。また、視点から画面までの距離を1とする。

【0106】この図に示す通り、P(X, Y)は画面上をP'(X', Y')へ移動するが、この間、この点は三次元空間においてS(x, y, z)からS(x', y', z')へと移動する。ここで、

$$(a - X'c)(Y' - Y) = (b - Y'c)(X' - X) \quad (式4)$$

式4は画面上の動き情報で表現されているため、工程1で得られた情報によって未知数a, b, cを決めることができる。しかしこの際、現実にはk倍の大きさの物体がk倍離れたところをk倍の速さで移動するケースにおいて、このkの値(スケールファクター)を決めることはできず、a, b, cについてはそれらの比のみを求めることが可能となる。数学的にいえば、(X, Y)と(X', Y')の対応を3組与えても、この連立方程式を行列表示した際の係数行列のランク(階数)は高々2であり、a, b, cは相対値としてしか決まらない。そこで本工程では、仮にc=1と正規化してa, bを表すことにする。比のみでも、次工程による処理が可能なのである。

【0108】並進運動の別の解法として、式4から誤差※

$$a \sum (Y' - Y)^2 - b \sum (X' - X)(Y' - Y) - c \sum (Y' - Y)(XY' - X'Y) = 0 \quad (式9)$$

$$-a \sum (X' - X)(Y' - Y) + b \sum (X' - X)^2 + c \sum (X' - X)(XY' - X'Y) = 0 \quad (式10)$$

$$-a \sum (Y' - Y)(XY' - X'Y) + b \sum (X' - X)(XY' - X'Y) + c \sum (XY' - X'Y)^2 = 0 \quad (式11)$$

以上が並進運動に関する計算方法の例である。

【0111】2. 動きが回転運動を含む場合

回転運動はx, y, z方向の3つの変位と各軸を中心とする3つの回転角、例えば α, β, γ によって記述することができる。回転角はオイラー角またはロールピッチ法などによって表現することができる。

【0112】ここで上記合計6つの変数を決定すればよいが、ここでも上述のごとくスケールファクターが決まらないため、ある変数を1として各変数の比を求める。理論上、代表点と対応点を5組とれば運動を記述することができる。

【0113】ここで注意すべきは、組のとりかたによっては動きの様子が線形解法によって求まらないことがある点である。こうした場合を考慮する際、組を8以上とればよいことが知られている。8組の変化から線形解法によって回転運動を記述しうる根拠については、例えば★

$$(x, y, z) = R^{-1} \{ (x', y', z') - (a, b, c) \} \quad (式12)$$

をカメラの動きと考える。

【0116】図13はカメラの三次元移動とある点Pの画面上の移動から点Pの三次元座標を導く原理を説明す

*【数4】

$$(x', y', z') = (x, y, z) + (a, b, c)$$

とすれば、画面までの距離が1なので、

$$X = x/z, \quad Y = y/z$$

$$X' = x'/z', \quad Y' = y'/z'$$

となる。これを解けば、

$$X' = (Xz + a) / (z + c)$$

$$Y' = (Yz + b) / (z + c)$$

となるため、zを消去し、次式が求められる。

【0107】

【数5】

※eを、

【数6】

$$e = \{ (a - X'c)(Y' - Y) - (b - Y'c)(X' - X) \}^2$$

$$= \{ (Y' - Y)a - (X' - X)b - (XY' - X'Y)c \}^2 \quad (式5)$$

と定義し、代表点と対応点の全対応関係についてeの総和 $\sum e$ をとり、この値を最小にするa, b, cを次の式から求めてもよい。

$$\frac{\partial (\sum e)}{\partial a} = 0 \quad (式6)$$

$$\frac{\partial (\sum e)}{\partial b} = 0 \quad (式7)$$

$$\frac{\partial (\sum e)}{\partial c} = 0 \quad (式8)$$

より具体的には、式6~8はそれぞれ次の形に展開される。

【0110】

【数7】

$$(式9)$$

$$(式10)$$

$$(式11)$$

★「動きからの単眼立体視による形状認識の線形解法について」(出口・秋場、計測自動制御学会論文集vol.26, No.6, 714/720 (1990))などに示されている。

【0114】[工程3] 奥行き情報の獲得

工程2によって映像各部位の三次元的な動きの相対量がわかった。工程3では、この相対量から各部位の奥行き情報を導出する。本工程では説明のために、被写体は静止しており、それを撮影するカメラの側が動くものと仮定する。映像処理の際には被写体とカメラの相対運動が問題となるため、この仮定によって良好な結果が得られる。

【0115】映像のある部位の動きを回転行列Rと並進ベクトル(a, b, c)により、

$$(x', y', z') = R(x, y, z) + (a, b, c)$$

と表す場合、この逆変換、

【数8】

る図である。同図からわかるように、この原理は一般に三角測量の原理として知られるもので、位置の異なる2点から点Pの方向を見たとき、点Pの現実の位置(図中

23

の点S)はそれら2つの視線の交点に存在するというものである。

【0117】同図では、時刻 $t \sim t'$ の間にカメラが矢印で示すように式12に従って移動したとする。フレーム t では点Sが点 P_t に、 t' では点 $P_{t'}$ にそれぞれ投影されている。点Sは図中の2つの直線 L_t 、 $L_{t'}$ の交点にある。

【0118】ここでカメラの方向と L_t 、 $L_{t'}$ のなす角 θ_t 、 $\theta_{t'}$ は既知であり、一方カメラの移動方向と距離が判明しているため、点Sの三次元座標を求めることが可能となる。この座標により、映像各部位の奥行き情報が判明する。

【0119】ここで注意すべきは、前述のごとく $c=1$ *

$$L_t : (x, y, z) + \alpha (u, v, w)$$

$$L_{t'} : (x', y', z') + \beta (u', v', w') \quad (式13)$$

と表すことができる。従って、

$$【数9】 e = \{(x+\beta u)-(x'+\alpha u')\}^2 + \{(y+\beta v)-(y'+\alpha v')\}^2 + \{(z+\beta w)-(z'+\alpha w')\}^2$$

とし、 e を最小にする α 、 β を $d e / d \alpha = 0$ 、 $d e / d \beta = 0$ より求める。すなわち、

$$【数10】 (u^2 + v^2 + w^2) \alpha - (uu' + vv' + ww') \beta + (x - x') u + (y - y') v + (z - z') w = 0$$

$$(u'^2 + v'^2 + w'^2) \beta - (uu' + vv' + ww') \alpha + (x - x') u' + (y - y') v' + (z - z') w' = 0$$

を解いて α 、 β を求め、最終的に点Sの奥行き値を、

$$【数11】 \{(z + \alpha w) + (z' + \beta w')\} / 2$$

とすればよい。ここで仮に誤差が0だったとすれば、この座標は両直線の交点の z 座標に一致する。

【0122】また別の方法として、これら両直線を一旦フレーム t の画面に透視投影し、投影後の最近接点の z 座標を求めることもできる。ここで L_t は代表点である一点に投影され、一方 $L_{t'}$ は一般に直線に投影され *

$$x = (m^2 X - kn - kmY) / (k^2 + m^2) \quad (式16)$$

$$y = (k^2 Y - mn - kmX) / (k^2 + m^2)$$

となる。ここで点Tに対応するもとの $L_{t'}$ 上の点をE (x', y', z') とすれば、点Eは、式16を式14に代入して β を求め、これを $L_{t'}$ の式に代入することによって求められる。ここで β は、

$$\beta = (xz' - fx') / (fu' - xw')$$

であるため、これを式13へ代入し、点Eの z 座標 z'' は、

$$z'' = z' + w' (xz' - fx') / (fu' - xw')$$

と求まる。これを点Sの奥行き値とすればよい。

【0125】なお、画像処理の際の誤差によって上記の奥行きが負の値となる場合(点Sがカメラの後方に存在することになる場合)、この計算結果は信頼することができない。このときは正の奥行き値を持つ近傍の代表点から補間するなどの処理を行う。

【0126】以上、いずれの方法をとるかは別として、求められた映像各部位の奥行きは、例えば代表点ごとに

24

*という正規化のため、求められた座標も一定の割合で拡大または圧縮されていることである。しかしこの場合でも、奥行き情報は一樣に拡大圧縮されているため、奥行きの相互関係は正しい。

【0120】以上が本工程の概要であるが、本工程では前工程までの誤差を考慮する必要がある。誤差により、通常は前記 L_t 、 $L_{t'}$ が計算上交わらないためである。こうした事情に配慮し、本工程では両直線の最接近点の midpoint の z 座標を点Sの奥行き値と近似する。これを数式によって説明する。

【0121】上記 L_t 、 $L_{t'}$ の方向ベクトルをそれぞれ (u, v, w) 、 (u', v', w') とする。ここで実数パラメータ α 、 β により両直線は、

※ $L_{t'}$ が式13で表記されるならば、投影後の各点の x 、 y 座標は、 $L_{t'}$ 上の各点のそれらをその z 座標で割ることにより、

$$x = f(x' + \beta u') / (z' + \beta w') \quad (式14)$$

$$y = f(y' + \beta v') / (z' + \beta w') \quad (式15)$$

と書くことができる。ここで f は視点からフレーム t の画面までの距離で、実際には $f=1$ などとして扱えばよい。式14、15から β を消去すれば投影後の直線(以下 L_i という)が以下のように求まる。

$$【0123】 kx + my + fn = 0$$

ただしここで、

$$k = v'z' - w'y', m = w'x' - u'z', n = u'y' - v'x'$$

【0124】求めるべき最近接点は、代表点 P_t から L_i に下ろした垂線と L_i の交点(以下Dとする)であり、その座標は、

数値として与えればよい。図14はフレーム t において各代表点に数値が与えられた状態を示す図で、例えば $P_t(2, 3)$ 、 $P_t(4, 3)$ それぞれの奥行きは100、200となり、後者の実際の位置は前者よりも2倍離れたところにあることがわかる。

【0127】[工程4]画像の生成

工程3で求められた奥行き情報に応じて視差を決定し、左右画像を生成する。このとき、奥行きが大きいほど視差を小さくする。

【0128】図15は奥行き情報に従って視差を与える様子を示す図である。同図では、被写体とカメラを含む系全体を上から見ている。ここでカメラによって撮影された映像画面に、図14で例示した $P_t(2, 3)$ 、 $P_t(4, 3)$ をとると、これらの実際の位置(同図ではそれぞれ $S_t(2, 3)$ 、 $S_t(4, 3)$ とする)はカメラから1:2の距離にある。

25

【0129】そこで本工程では、まずLスクリーンとRスクリーンを図の位置に置く。次に人の左右両眼に対応するL視点、R視点を設定し、これら各視点からSt(2,3)、St(4,3)を見て前記L、Rスクリーンに投影する。この作業をすべての代表点について行い、最終的にL、Rスクリーンに得られた画像をそれぞれ左眼用、右眼用画像とする。以上が本工程の概要であり、こうして得られた両眼用の画像を例えば特開平3-65943号公報に記載のレンチキュラー方式のディスプレイに表示すれば、良好な立体視が実現される。

【0130】本実施形態では、所期の領域を切り出した上で立体表示を行ってもよい。例えば奥行き5メートルの位置に「人」がいて、背景に「山」がある映像の場合、「奥行き10メートル以内」という条件を与えて画像処理を行うことにより、「人」の領域のみを切り出すことができる。そこで「人」の領域のみに対して左右画像を生成し、それ以外の領域を空白にしたり、予め用意した別の映像を貼り付けることができる。本発明はこうした画像認識・加工技術をも与える。

【0131】本工程といままでの工程との違いの1つに、使用する映像フレームの数がある。いままでの工程では少なくとも2つの映像フレームから必要情報を抽出したが、本工程では1つの映像フレームから左右画像を生成することができる。図16はフレームtを基準として生成された左右画像を示す図で、フレームtの映像各部位「木」、「家」、「人」はこの順に奥行きが小さくなる。最も近い「人」には、

1. 右眼画像において最も左に変位し、
2. 左眼画像において最も右に変位する、
という特徴がある。1. はもとのフレームtの実際の風景を少し右に移動して見た状態、2. はその逆と考えればよい。この結果、「人」の奥行きは小さくなり、近い距離に知覚される。同図では各部位の変位がメッシュ格子点の移動によって表されており、「人」、「家」、「木」の順に変位(視差)が小さくなる。

【0132】フレームtから画像を生成する場合、図16でメッシュ分割された領域を変形させることになるが、この領域変換の際、条件に応じて以下のように線形または非線形変換を使い分ける必要がある。

【0133】1. 非線形変換

図16からわかるように、領域のいくつかは台形に変換されるべきであるが、アフィン変換等の通常の線形変換でこれを実現することはできない。そこで領域の4頂点に対する変換を行う場合は、透視変換等の非線形変換を用いる。

【0134】2. 線形変換

領域を2つの三角形に分割するのであれば、その3頂点に対してアフィン変換等の線形変換を採用することができる。

【0135】以上の変換によれば、画像各部位が水平方

26

向に変位する結果、画像の枠部に段差が生じることがある。例えば図16において、左右画像がともに画像内側に変位している箇所が該当する。本工程では、このような箇所に画素を追加して段差を修正し、画像領域の形状(この場合は矩形)をもとのまま維持することにする。

【0136】不足画素の奥行きは、例えばその近傍画素の奥行きを補間する方法によって求めることができる。こうして追加された画素は両眼のうち一方からのみ見えることになるが、これは人が実際に窓から外を見る場合にも窓枠付近で発生する現象であり、不自然ではない。なお当然ながら、段差によって不足した画素を追加するほかに、段差によって外側にはみ出した余剰画素を削除して形状を維持してもよい。いずれにせよ、こうした処理によって不自然な表示を回避することができる。

【0137】以上、本工程では奥行きに従って視差を決定したが、以下の理由からいったん決まった視差について調節可能な構成とすることが望ましい。

【0138】1. 眼の疲れの考慮

上記の例でいえば、最も近い「人」であってもその奥行きが極端に小さいことは好ましくない。画面よりも極端に前に像が結ばれると眼が疲れやすいことが知られているためである。別の報告(日経エレクトロニクス1988年4月4日号211ページに記載)によれば、目から50cmの距離にディスプレイを置いたとき、画像各部位の奥行きが0.2~2m程度の範囲にあることが望ましいとされる。

【0139】2. 個人差の考慮

視差の広がりや大ききとった画像(近い物)はより手前に、遠い物はより遠くに見える画像を好む人もいれば、その逆もある。

【0140】3. 処理能力の考慮

例えば山など遠景を構成する領域をすべて一定の奥行きとして画像を生成すれば処理すべきデータ量が減る。

【0141】以上の理由から、本工程では必要に応じて以下の奥行き変換または視差変換機能を採用する。

【0142】1. 奥行き変換機能

奥行きを直接線形または非線形変換する。変換の対象は奥行きであり、ここでは視差は結果的に変更されることになる。例えば奥行きが $a \sim 10a$ (a は任意)の間をとる映像については、これを一様に10倍し、 $10a \sim 100a$ とすることができる。これは奥行きが全体的に小さ過ぎる場合に有効な措置である。

【0143】また別の例として、奥行きが $0 \sim 100a$ のとき、これを任意の点、例えば $50a$ を中心として $25a \sim 75a$ へ圧縮変換してもよい。さらに他の例として、 $20a$ 以下の奥行きはすべて $20a$ に、一方 $1000a$ 以上はすべて $1000a$ に変換する方法も考えられる。ただしこの場合、上下限値である $1000a$ および $20a$ において変換結果が不連続となるため、映像によ

27

っては多少不自然な表示となる。この状態を解決する方法としては、前記上下限值に滑らかに収束するような非線形変換、この場合であればもとの奥行きを z として、
 $z \rightarrow \alpha / |1 + \exp(-(x - 0.5 \alpha) / \alpha T)| + z_0$
 ただし、

$$z_0 = 20 \alpha, \alpha = 1000 \alpha - 20 \alpha = 980 \alpha, T = 4$$

のような変換を行えばよい。

【0144】2. 視差変換機能

視差を線形または非線形変換する。奥行きからいったん視差を導出し、この視差に対して変換を加え、これを再度奥行きに戻す方法である。

【0145】図17は視差を非線形変換する様子を示している。同図の中心線L上に対象となる点Sをとり、一方、視点AからLに下ろした垂線の足を点Bとする。このとき、点Sの奥行きは線分SB、点Sの視差 θ （厳密に定義に従えば視差の $1/2$ ）は図のように与えられる。

【0146】ここで、一例として視差を $1/2$ 倍に一樣圧縮することを考える。すなわち、

$$\theta' = \theta / 2 \quad (\text{式17})$$

となる点S'が変換後の点であり、この奥行きは線分S'Bによって与えられる。この変換に関する一連の処理をより数学的に説明すれば、まず奥行きSBから θ を、

$$\theta = \text{atan}(SB)$$

によって求め、S'Bを、

$$S'B = \tan \theta'$$

によって求め、このS'Bを変換後の奥行き情報として利用すればよい。この変換は奥行きの単純な線形変換よりも、遠い点はより遠くへ、近い点はより近くへ変換されるため、奥行き感の調節効果が大きい。なお式17は単純な線形圧縮式であるが、 $\theta \rightarrow \theta'$ についても1.で示した各種非線形変換を施すことができる。

【0147】以上が本発明の実施形態1に係る三次元表示画像生成方法の説明である。本発明によれば、既存の映像の組合せではなく、画像を奥行き情報に従って新たに生成するため、従来技術の技術的・機能的課題を解決することができる。この際、時間差方式の前提条件である被写体の水平方向の動きは不要なため、本発明の適用範囲は広い。また、代表点に対する対応点の検出方法を開示したため、奥行き情報の自動抽出、画像の自動生成を容易かつ効果的に行うことができる。

【0148】**実施形態2.** 実施形態2では実施形態1の実施に最適な装置を説明する。

【0149】図18は実施形態1を実施するためのハードウェア構成例を示す図である。

【0150】同図において、処理対象となる映像は映像入力回路20から入力され、ここでA/D変換を受ける。変換後の映像はフレームメモリ制御回路22によ

28

てフレームメモリ24に格納される。フレームメモリ24の後段には、複数の映像フレームを読み出して対応点を検出する対応点検出回路26が置かれている。この回路は実施形態1の工程1をハードウェア的に実施するもので、ブロックマッチングには例えばMPEGのエンコード回路を採用することができる。

【0151】対応点検出回路26において検出された対応点の座標は、いったん対応点座標メモリ28に格納され、動き検出回路30によって適宜読み出される。動き検出回路30は実施形態1の工程2、3を実施するもので、並進・回転運動から被写体の三次元相対位置を算出する。

【0152】つづいて三次元位置情報は画像生成回路32へ入力される。この回路はフレームメモリ24から生成の元映像を読み出し、所定の視差を与えて左右画像を生成する。画像生成回路32の前段には、外部から各種指示を受け付ける指示入力部34が設けられている。

【0153】こうして画像生成回路32によって生成された左右画像は、画像出力回路によるD/A変換を受け、図示しないディスプレイ装置へ出力される。

【0154】以上の構成による本装置の動作を説明する。

【0155】まず、カメラによって撮影された映像またはビデオ再生装置によって再生された映像等が映像入力回路20から入力される。この映像はフレームメモリ24に格納されるが、本装置において三次元表示ではなく通常の表示を行う場合には、フレームメモリ24に格納された映像を順次読み出してディスプレイに表示するか、または入力された映像をそのままディスプレイに表示することで対応する。三次元表示を行う場合は、フレームメモリ24から読み出された複数の映像フレームに対し、対応点検出回路26、動き検出回路30によって被写体の奥行き情報が獲得される。つづいて画像生成回路32がこの奥行き情報に従い、左右画像を生成する。

【0156】以上が本装置の概要である。なお、前述の指示入力部34の構成と機能は以下のように実現できる。

【0157】1. 指示入力部34が調節つまみの場合
 ユーザーの好みに応じて、生成された画像の奥行き感を調節する。このつまみによって奥行き感を圧縮または拡大し、奥行き感を可変とする。奥行き感を最小に設定した場合に通常の二次元表示に戻るよう、予めつまみの回転量を調整しておく方法もある。

【0158】2. 指示入力部34がポインティング・デバイスの場合

(1) 奥行き感の調節を画像全体ではなく、画像各領域毎に行う。例えば図16において「人」をより近い位置に表示したい場合、マウス等のポインティング・デバイスで「人」をクリックする。この結果、画像生成回路32は「人」の奥行き情報を内部で変換し、視差を大きく

29

して強調表示することができる。この際、奥行き感の変更に応じてその領域の表示面積を変更すればさらに効果的である。より具体的には、奥行きを1/2倍したらその領域の面積を4倍する等の方法がある。

【0159】(2) 視点変更映像を得ることができる。映像の撮影視点をマウス等でクリックして指示すれば、奥行き情報が利用できるため、視点の変更に伴って発生すべき映像各部位の並進・回転運動を逆算することができる。このため、視点変更後に見えるべき映像を再構築することが可能となる。例えば図16において、カメラ

の設置される高さを変えた場合、カメラを前進、後退させた場合の映像を生成することが可能となる。この際、再構築された映像について映像各部位の奥行き情報を再計算し、新たな奥行き情報から視差を補正すれば三次元表示も良好に保たれる。なお、視点変更映像については実施形態4でも記述する。

【0160】以上、本装置をワークステーションに組み込んで実験した結果を図面によって説明する。

【0161】図19～26は本装置による画像生成の過程を示すもので、いずれもディスプレイ上に表示された中間調画像の写真であり、約640×480画素領域で構成されている。

【0162】図19と図20はそれぞれフレームtとフレームt'の映像で、カメラの移動によって多少映像に動きがある。図21はフレームtをメッシュ分割して代表点を設けた状態を示し、図22はフレームt'における対応点の初期位置を示している。ここでは特徴点を優先的に扱い、代表点を中心とする16×16画素の領域でブロックマッチングを行った結果の最良点を初期位置とした。

【0163】図23は実施形態1の式3、つまり対応点の位置関係をも考慮して対応点位置を改善した結果を示す図で、図22に比べて位置の大幅な改善が見られる。

【0164】図24は獲得された奥行き情報を濃淡で表したもので、薄い画素ほど奥行きが小さい。同図からわかるように、奥行き情報がかなり正確に得られている。

【0165】図25と図26はそれぞれ、奥行き情報に従って生成された左右画像である。これらの図からわかるように、手前の物ほど視差が大きくなり、与えられる水平変位も大きくなっている。

【0166】以上、本装置によれば実施形態1の方法を円滑に実施することができる。この際、特にブロックマッチングのソフトウェアによる実行時間を考慮すれば、ここをハードウェアで行うことによる処理速度の改善効果は大きい。

【0167】本装置を製品とする場合の態様としては、例えばパーソナルコンピュータやワークステーションに図18の構成を有するアド・オンカードを搭載する方法、テレビ受信装置やビデオ再生装置に図18の構成を予め内蔵する方法等が有効である。本装置をカメラと組

30

み合わせた場合、例えば物のまわりを一周して映像を取り込み、これから三次元表示カタログを作成するなどの応用もある。この方法によれば、従来必要とされたレーザ、赤外線、超音波等による奥行きの測定が不要となる。

【0168】実施形態3。実施形態1、2では、入力映像が単眼カメラによって撮影されるものとした。ここでは、多眼カメラによるステレオ映像を入力映像とする場合の三次元表示画像生成方法を、実施形態1との相違点を中心に説明する。

【0169】図27は実施形態3によって三次元表示画像を生成するための主な工程を示す。同図と実施形態1の図4との主な違いは以下の通りである。

【0170】1. 工程1の「動き情報」が「変位情報」に変更される

実施形態1では異時刻フレームを扱ったが、実施形態3では基本的に同時刻フレームを扱う。同時刻の場合、被写体に動きを定義することはできないため、代わりに同時刻フレーム間の被写体の位置のずれ(変位)の情報を抽出する。

【0171】2. 工程2が不要となる

図4の工程2「三次元動き情報の算出」に対応するステップがない。多眼の場合、はじめから図13の状態では撮影がなされるため、三角測量の原理によって直接奥行き情報が獲得できるためである。

【0172】なお、複数カメラの相対位置関係に狂いが発生しうる多眼カメラシステムを用いる場合、この狂いを補正するセルフキャリブレーションを行ったほうがよい。この場合、工程2はセルフキャリブレーション工程となる。セルフキャリブレーションの手法については、例えば、富田、高橋「ステレオカメラのセルフキャリブレーション」(情報処理Vol.31, No.5 (1990) 650～659ページ)、特開昭63-293038号公報、特開昭63-293039号公報などに示されている。

【0173】以下、実施形態3の工程1～3を説明する。

【0174】[工程1] 二次元変位情報の抽出

実施形態1の説明において、「動き」を「変位」に置き換える他、フレームt、t'の組をフレーム1、2に置き換えればよい。フレーム1、2はそれぞれカメラ1、2から撮影された映像を指し、撮影時刻はtで固定とする。実施形態3では、最低これら2枚のフレームのみから最終画像を得ることができる。すなわち多眼撮影の場合は、入力静止映像であってもよい。その他、実施形態1の工程1との相違は以下の通りである。

【0175】(1) 実施形態1のS11(対応点候補領域の設定)では、映像の動きの激しさまたは各部位の移動軌跡に基づき、異時刻フレームの選択または対応点候補領域を絞り込んで対応点検出処理の計算量削減を行った。実施形態3では絞り込みの方法を以下のように変更

31

し、同様に有効な計算量削減を実現する。

【0176】まず、通常どおり多眼カメラが水平に設置されると仮定する。このとき、対応点のy座標（上下方向の座標）はほぼ等くなる。この仮定と画像処理に伴う誤差およびカメラの設置誤差を考慮し、対応点候補領域を横長の帯状領域に限定する。さらに、フレーム t'

（ $t' = t - 1$ ）において対応する代表点の位置の差が x であれば、フレーム t における対応点探索領域も、やはり差が x となる近傍に限定することができる。

【0177】（2）実施形態1のS12（対応点候補領域における非類似度の計算）では、映像の動きが緩慢過ぎる場合に統計処理を導入したが、実施形態3ではこの作業も不要である。

【0178】（3）実施形態1のS12同様、実施形態3でも対応点の位置を決めるべくブロックマッチングを行うが、ここではバイアスド・ブロックマッチングを採用したほうがよい場合がある。バイアスド・ブロックマ

$$E1 = \sum \sum \{ I1(P1x+u, P1y+v) - I2(P2x+u, P2y+v) \}^2 \quad (式18)$$

と簡単になる。この式が濃淡画像の場合の通常のブロックマッチングを表す。

【0180】一方、バイアスド・ブロックマッチングで

$$E1 = \sum \sum \{ I1(P1x+u, P1y+v) - I2(P2x+u, P2y+v) - \alpha \}^2 \quad (式19)$$

とすればよい。カラー画像の場合、 α は α_R 、 α_G または α_B であり、RGBそれぞれの画像において求めたE1の和、すなわち $E1_R + E1_G + E1_B$ でマッチングを行う。さらに見やすさを考えて、 $I1(P1x+u, P1y+v)$ を単に $I1$ 、 $I2(P2x+u, P2y+v)$ を単に $I2$ と表記すれば、式19は、

$$E1 = \sum \sum (I1 - I2 - \alpha)^2 \quad (式20)$$

となる。 $I1$ 、 $I2$ は u 、 v の関数であるが、 α は定数★³⁰

$$\begin{aligned} E1 &= \sum \sum \{ (I1 - I2)^2 - 2\alpha(I1 - I2) + \alpha^2 \} \\ &= \sum \sum (I1 - I2)^2 - 2\alpha \sum \sum (I1 - I2) + \sum \sum \alpha^2 \end{aligned} \quad (式21)$$

領域の総画素数を N と仮定すれば、 $\sum \sum 1 = N$ だから、式21は、

$$E1 = \sum \sum (I1 - I2)^2 - 2\alpha \sum \sum (I1 - I2) + N\alpha^2 \quad (式22)$$

となる。従って、

$$dE1/d\alpha = -2\sum \sum (I1 - I2) + 2N\alpha$$

であるから、

$$\alpha = \{ \sum \sum (I1 - I2) \} / N \quad (式23)$$

のとき、 $E1$ は最小となる。この α は、ブロックマッ

$$E1 = \sum \sum (I1 - I2)^2 - \{ \sum \sum (I1 - I2) \}^2 / N \quad (式24)$$

となるため、結局バイアスド・ブロックマッチングでは式24を計算すればよいことになる。あとは実施形態1同様の処理を経て最良マッチングを探索すればよい。

【0182】ここでは当然ながら、HVC濃度などRGB濃度以外の色空間による濃度を採用してもよい。2乗誤差の代わりに1乗誤差、すなわち残差に基づいてブロックマッチングを行ってもよい。

【0183】（4）実施形態1のS13（対応点の初期位置の決定）では、特徴点として、異時刻フレーム t 、

32

*ッチングは、多眼カメラを構成する各カメラが異なる特性を持つ場合、有効に機能する。例えば、カメラ2がカメラ1よりも青味がかかった映像を送出するとすれば、フレーム2の色濃度から青（B）の成分を一定量差し引いた後（すなわち色偏向定数 α_B を引いた後）、ブロックマッチングを行うべきである。こうした処理を行わない場合、最良マッチングを取り逃がすおそれがある。実際には、例えばRGBで色濃度を表す場合、前記の青（B）のみならず、赤（R）と緑（G）についてもそれぞれ色偏向定数 α_R 、 α_G を引くべきである。

【0179】図7および式1を基礎として、バイアスド・ブロックマッチングを式で説明する。ここでは、実施形態1で用いた $Pt(i, j)$ をフレーム1、2に対応して単に $P1$ 、 $P2$ と表記し、 $I1(i, j)$ も同様に $I1$ 、 $I2$ と表記する。このとき式1は、

【数12】

※は、式18を、
20 【数13】

★である。

【0181】 α の最適値を考える。カメラ1、2はほぼ同じ被写体を撮影しているはずだから、フレーム1、2の映像は、映像各部位の変位は別として、ほぼ同様の内容を含む。すなわちカメラの特性が近づくほど、式20の $E1$ の値は小さくなる。この事実から逆に、 α は $E1$ を最小にする値とすべきことがわかる。式20は、

【数14】

☆【数15】

◆ングの対象となる2つの領域間の各画素の色濃度差の平均値と言い換えられる。式23を式22に代入して計算すれば、

【数16】

t' 、・・・において位置が安定的に変化している点を選んだが、ここでは選定基準を加重する。

【0184】図28は実施形態3で導入される特徴点の選定基準を示す図である。同図において、F10～F12の3枚がカメラ1によって撮影された異時刻フレーム、F20～F22の3枚がカメラ2によって撮影された異時刻フレームである。それぞれ左右2枚の組が同時刻フレームを示す。ここではある点Pに注目し、点Pの位置の異時刻フレーム間の動きをベクトル A_n で、点P

33

の同時刻フレーム間の変位をベクトル B_n (n : 自然数) でそれぞれ表している。

【0185】以上の設定の下、実施形態3では、次の基準を満足する点を特徴点として選定する。

【0186】(a) ベクトル B_n がほぼ一定であるか、ほぼ一定変化する

あるいは、この他、

(b) ベクトル A_n がほぼ一定であるか、ほぼ一定変化する

の基準も追加し、(a)、(b) をともに満足する点を特徴点として選定してもよい。

【0187】(b) は実施形態1で導入した条件に相当する。既述のごとく、多眼撮影では同時刻フレームのみから奥行き情報を求めることができる。しかしその前提となる映像間の対応関係の正確な把握は、これとは別問題であり、異時刻フレーム間の情報も積極的に利用すべきである。上記の2条件を同時に満たす点は、相当正確に追跡されていると考えられるため、二次元変位情報の抽出に対して重要な手掛かりを与える。ただし入力が静止映像の場合は、既知の動的計画法(ダイナミック・プログラミング)によって対応点を求めることもできる。

【0188】[工程2] 奥行き情報の獲得

工程1で求められた映像各部位の変位から各部位の奥行き情報を導出する。多眼の場合、ある時刻 t で図13の状態が実現されるので、以下、実施形態1の工程3の方法によって奥行き情報を獲得すればよい。

【0189】ここで注意すべきは、撮影カメラの位置関係が固定であるため、この関係およびカメラの倍率(焦点距離)が既知であれば、実施形態1で決まらなとされたスケールファクター c も含め、実施形態3では正しい奥行き情報が求まる。

【0190】[工程3] 画像の生成

実施形態1の工程4(画像の生成)と同等の処理を行えばよい。

【0191】以上が実施形態3の概要である。実施形態3の場合、入力がステレオ映像、出力が三次元表示画像である。従って、入力映像の忠実な再現が可能となる他、実施形態1で説明した強調表示等の処理を施すことによって、所期の画像を生成することができる。

【0192】実施形態4。実施形態2では、マウスクリックによって視点変更映像を得る技術を説明した。ここでは、異なる目的によって視点変更映像を生成し、これを異なる用途に利用する例を説明する。

【0193】前述のように、本発明によれば、実際にカメラの位置を移動しなくとも、異なる視点からの映像を得ることができるが、この場合、当然ながら視点の仮想的な移動距離が短いほうが正確な画像を得ることができる。この事実を利用して、実施形態4では次の用途を考える。

【0194】1. 2眼映像から多視点映像の生成

34

2眼カメラによってステレオ映像が得られているとき、仮想的に第3のカメラを設けて多視点映像を生成する。2眼カメラの間隔は一般に小さいため、例えばこれらとともに、小さな三角形をなす点を視点とする視点変更映像を生成する。この映像は比較的正確であり、もとの2眼カメラによる映像と合わせることで良好な多視点映像を形成することができる。この後、奥行き情報を加味することによって、多視点のうちの任意の視点からの三次元表示画像を生成することができる。

【0195】2. スローモーション映像の生成

例えば時間的に最も近い2枚の異時刻フレームをフレーム t 、 t' とする。これらのフレームを撮影した視点をそれぞれ視点 t 、 t' とする。視点はフレーム t 、 t' の間に視点 t から視点 t' に実際に移動しているが、この間の映像は存在しない。そこで視点 t 、 t' 間に仮想的な視点を設け、視点変更映像を生成する。視点変更映像を多数設けて順次表示すればスローモーション映像が得られる。この際、次の効果がある。

【0196】a. 本来ならコマ送りの映像となるべき映像の動きが極めて滑らかになる

b. 時間的に近いフレーム間では一般に視点の移動が少ないため、非常に良好な映像が得られる

c. 視点 t 、 t' 間の視点の移動経路を変化させれば、異なるイメージのスローモーション映像が得られる。

【0197】この間も、奥行き情報を加味することによって滑らかな三次元表示画像を生成することが可能となる。なおこの技術は、当然ながら同時刻フレーム間に適用することもできる。

【0198】

【発明の効果】本発明によれば元の映像の奥行き情報から三次元表示用の画像を生成することができ、課題であった機能的・技術的制約が解消される。入力映像は単眼、多眼のいずれでもよく、適用範囲が広い。単眼の場合はシステム構成が簡素化でき、多眼の場合は奥行きの計算精度がさらに高くなる。

【0199】本発明によれば、複数の映像フレーム間における映像各部位の二次元的な位置の変位から奥行き情報を抽出することができ、容易に三次元表示を行うことができる。

【0200】二次元的な位置の変位量に基づいて複数の映像フレームを選択する場合は、計算途上の誤差が減る。

【0201】二次元的な位置の変位量を統計的に処理する場合は、選択の最適化、客観化が可能となる。

【0202】映像各部位の動きベクトルの分散が所定値よりも大きくなるように複数の映像フレームを選択する場合は、選択の客観化に加え、計算精度の保証が可能となる。

【0203】分散が大きくなるような選択ができない場合、通常の映像表示等、他の表示で代替できる。

【0204】二次元的な位置の変化から映像各部位が現実の三次元空間において占める相対位置関係を導出する場合は、奥行きを正しく求めることができる。

【0205】映像各部位の三次元的な動きから三角測量の原理によって映像各部位の三次元空間における位置座標を算出する場合は、奥行きの決定を容易かつ正確に行うことができる。

【0206】本発明では、代表点と対応点という概念を導入したため、従来のように水平方向の動きに限らず、三次元表示が可能となる。

【0207】代表点と対応点の位置関係から他の異時刻フレームの対応点の位置を予想する場合は、対応点探索の計算量が削減される。

【0208】代表点のうち幾何的に特徴のある映像領域に関連するものについては、対応点の関連する映像領域についてもその特徴が維持されるため、対応点の位置精度が高い。

【0209】幾何的に特徴のある映像領域として直線を含む領域を採用する場合は、対応点の位置精度が改善される。

【0210】複数の映像フレーム間で映像の類似度が高い点を対応点とする場合は、対応点を正しく求めることができる。

【0211】映像の類似性のみならず、対応点間の相対位置の妥当性も評価する場合は、両評価の結果から対応点の位置精度を改善することができる。

【0212】前記各評価の結果が統合され、繰り返し計算がなされる場合は、対応点の位置精度を徐々に改善していくことができる。

【0213】前記類似性の評価をブロックマッチングによって行う場合は、容易かつ妥当な評価が可能となる。

【0214】ブロックマッチングでは、色濃度の n 乗誤差が計算されるため、映像に応じて実験等により最適の条件を設定することができる。

【0215】バイアスド・ブロックマッチングを行う場合は、特に多眼撮影された入力映像に対して良好な三次元表示画像の生成が可能となる。

【0216】色偏向定数を n 乗誤差総和が最小となるよう決められる場合は、評価結果の信頼性が高まる。

【0217】色偏向定数として映像領域間の各画素の色濃度差の平均値を採用する場合は、2乗誤差総和が最小となり、評価結果の信頼性が高まる。

【0218】特徴点の対応点を優先的に決める場合は、対応点の位置精度が高まる。

【0219】特徴点である対応点を補間することで非特徴点である対応点を決める場合は、非特徴点の対応点の位置精度も必然的に高くなる。

【0220】複数の映像フレーム間でブロックマッチングを行って特徴点を決める場合は、特徴点選定の客観化、選定結果の妥当性向上が可能となる。

【0221】異時刻フレームで位置の変化が安定して発生する点を特徴点とした場合、かかる点は精度よく追跡されているため、その位置精度が高い。

【0222】同時刻フレーム間における位置の変位が、その近傍時刻に撮影された同時刻フレーム間においてもほぼ一定となる点を特徴点とする場合、多眼映像から三次元表示画像を良好に生成することができる。

【0223】本発明によれば、奥行き情報に従った両眼視差から左右画像が生成されたため、現実の奥行き間と生成された画像の対応関係が良好となる。

【0224】映像各部位ごとに両眼視差を決定する場合は、生成された画像がより自然なものとなる。

【0225】所定範囲の奥行きを持つ映像部位のみを選定して左右画像を生成する場合は、一種の画像認識に基づく画像加工が可能となる。

【0226】画像に対して両眼視差に応じた水平方向の変位を与える場合は、人の眼にとって自然な三次元表示が実現される。

【0227】前記水平方向の変位が与えられたとき、画像枠に現れる段差を修正する場合は、この部分に不自然な表示がなされない。

【0228】奥行きが大きくなるに従って両眼視差を小さな値にする場合は、現実の映像を見た場合と同じ関係が実現される。

【0229】奥行きあるいは両眼視差を所定の範囲に収める場合は、眼の疲労等に配慮することができる。

【0230】奥行きまたは両眼視差の広がり任意の点を中心として線形圧縮する場合は、単純な変換で所望の奥行き感を得ることができる。

【0231】所定の範囲に収まらない奥行きまたは両眼視差をその範囲の上限値または下限値のうち近いほうの値に一律変換する場合は、容易に効果的な変換をなしつつ、画像データの圧縮、処理時間の短縮を図ることができる。

【0232】変換後の値が所定の範囲の上限値または下限値に滑らかに収束する非線形変換を用いる場合は、これら上下限値付近において表示が不自然とならない。

【0233】本発明によれば、生成されるべき三次元表示画像における所定の領域の奥行き感を変更することができ、見る者の好み等に応じた柔軟な対応が可能となる。

【0234】奥行き感の変更に応じてその領域の表示面積を変更する場合は、例えば近くに表示された物は大きくなり、自然な表示が可能となる。

【0235】仮想的な視点の変更に伴って映像を再構築する場合は、各種画像処理への途が開かれる。一旦多視点映像またはスローモーション映像を生成した後に三次元表示画像を生成する技術が可能となる。

【0236】このとき、再構築された映像について奥行き情報が再計算されるため、再度良好な三次元表示をす

37

ることが可能となる。

【図面の簡単な説明】

【図1】 背景は変化せず被写体のみが左から右へ移動する映像を示す図である。

【図2】 再生された右眼画像と左眼画像の間に一定の時間差を設ける様子を示す図である。

【図3】 図2の時間差の結果、視差が生じる様子を示す図である。

【図4】 実施形態1によって三次元表示画像を生成するための主な工程を示す図である。

【図5】 映像フレーム間の対応関係を検出するためのフローチャートである。

【図6】 基準フレーム t に代表点を設定する様子を示す図である。

【図7】 ブロックマッチングの様子を示す図である。

【図8】 仮の対応点 $Pt'(i, j)$ ごとに E の値を縦方向に示した模式図である。

【図9】 S12ステップの結果求められた対応点と代表点の関係を示す図である。

【図10】 対応点の相対位置を評価する原理を説明する図である。

【図11】 図9の対応点候補に対して本ステップの改善処理を行った結果を示す図である。

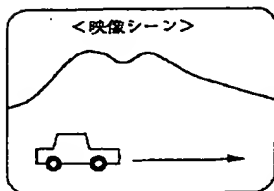
【図12】 ある点 P の画面上の移動と三次元空間での移動の対応を示す図である。

【図13】 カメラの三次元移動とある点 P の画面上の移動から点 P の三次元座標を導く原理を説明する図である。

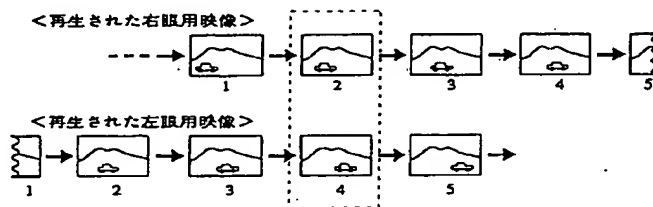
【図14】 フレーム t において各代表点に数値が与えられた状態を示す図である。

【図15】 奥行き情報に従って視差を与える様子を示す図である。

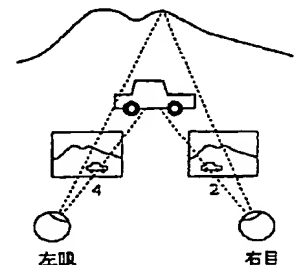
【図1】



【図2】



【図3】



38

* 【図16】 フレーム t を基準として生成された左右両画像を示す図である。

【図17】 視差を非線形変換する様子を示す図である。

【図18】 実施形態1を実施するためのハードウェア構成例を示す図である。

【図19】 フレーム t の映像をディスプレイ上に表示した中間調画像の写真である。

【図20】 フレーム t' の映像をディスプレイ上に表示した中間調画像の写真である。

【図21】 フレーム t をメッシュ状に分割して代表点を設けた状態をディスプレイ上に表示した中間調画像の写真である。

【図22】 フレーム t' における対応点の初期位置をディスプレイ上に表示した中間調画像の写真である。

【図23】 対応点位置を改善した結果をディスプレイ上に表示した中間調画像の写真である。

【図24】 奥行き情報を濃淡でディスプレイ上に表示した中間調画像の写真である。

【図25】 奥行き情報に従って生成された左画像をディスプレイ上に表示した中間調画像の写真である。

【図26】 奥行き情報に従って生成された右画像をディスプレイ上に表示した中間調画像の写真である。

【図27】 実施形態3によって三次元表示画像を生成するための主な工程を示す図である。

【図28】 実施形態3で導入される特徴点の選定基準を示す図である。

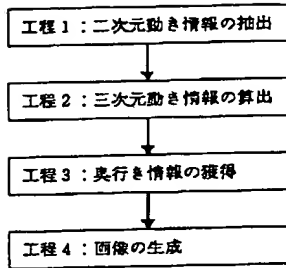
【符号の説明】

20 映像入力回路、22 フレームメモリ制御回路、24 フレームメモリ、26 対応点検出回路、28 対応点座標メモリ、30 動き検出回路、32 画像生成回路、34 指示入力部、36 画像出力回路。

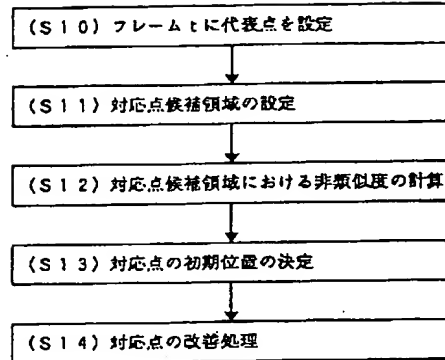
30

*

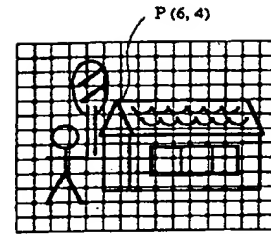
【図 4】



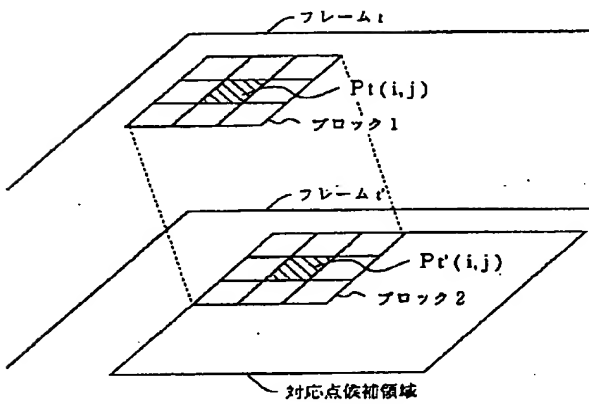
【図 5】



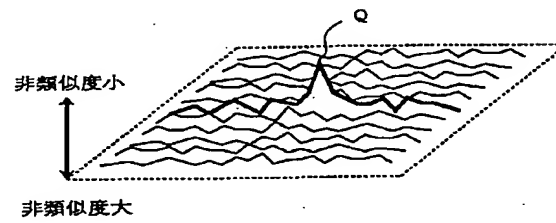
【図 6】



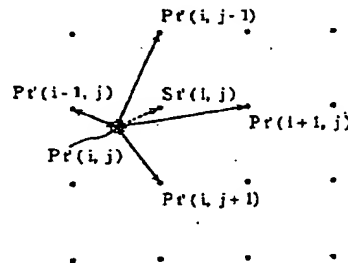
【図 7】



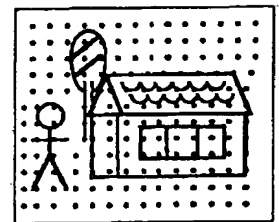
【図 8】



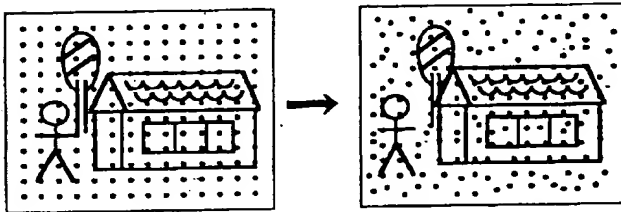
【図 10】



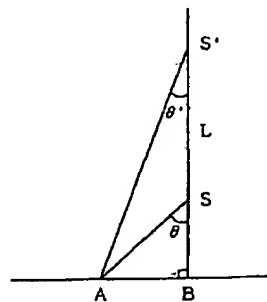
【図 11】



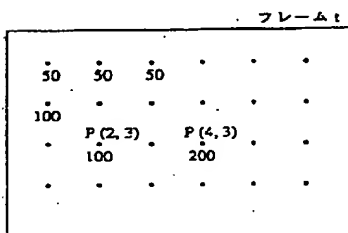
【図 9】



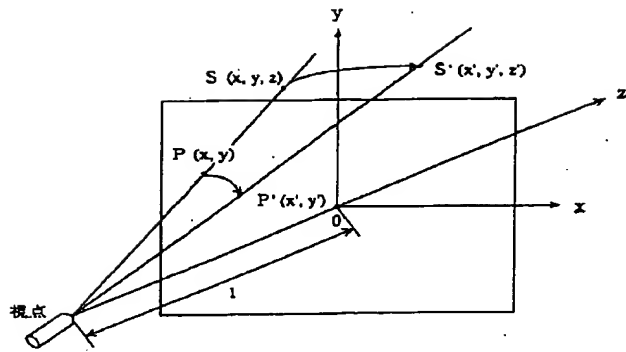
【図 17】



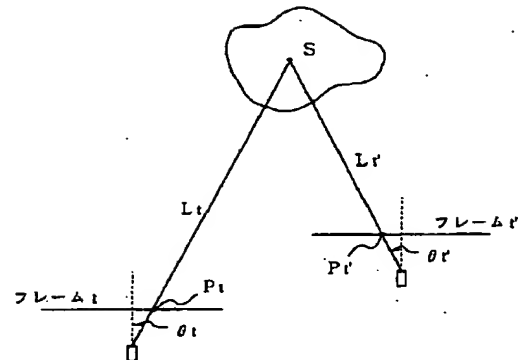
【図 14】



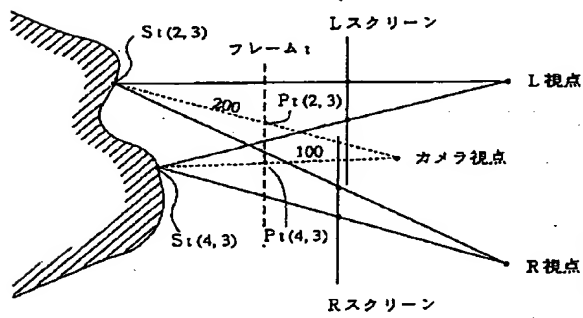
【図 12】



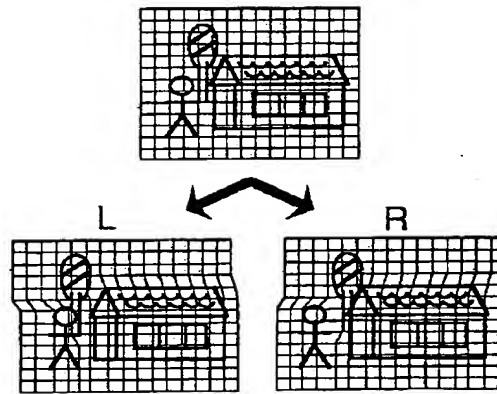
【図 13】



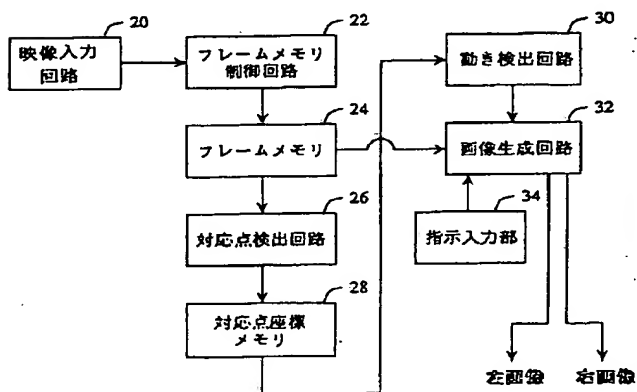
【図 15】



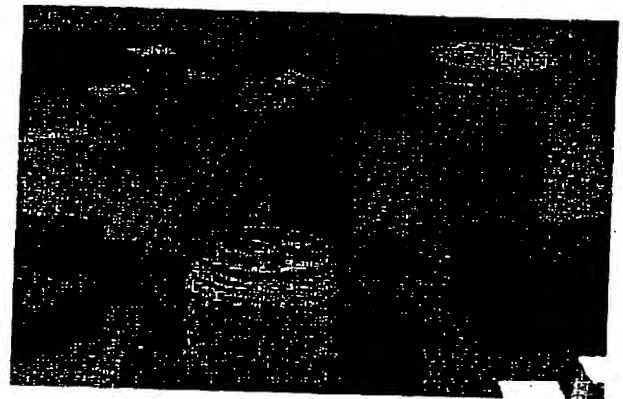
【図 16】



【図 18】



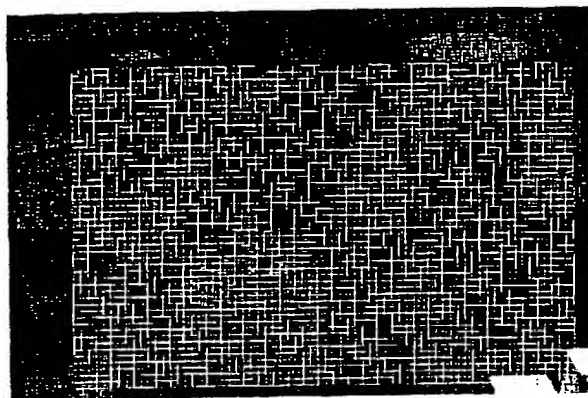
【図 19】



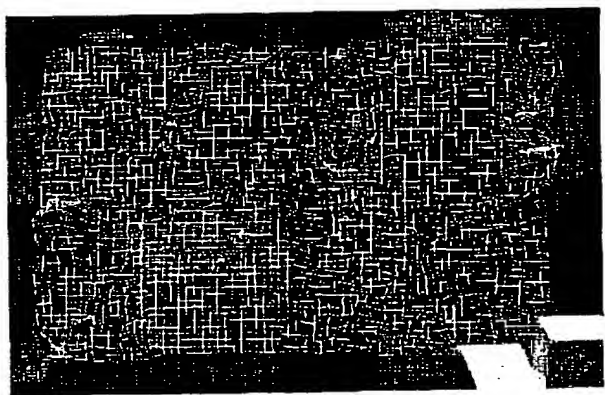
【図 2 0】



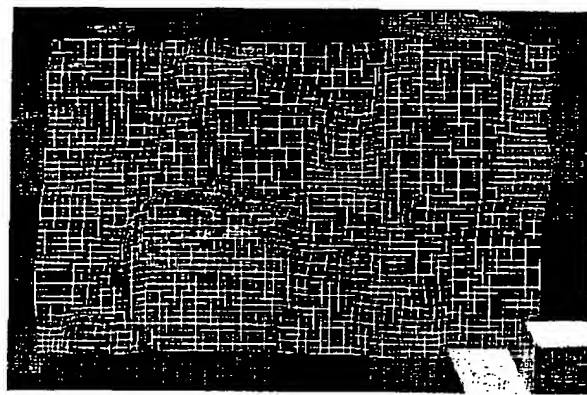
【図 2 1】



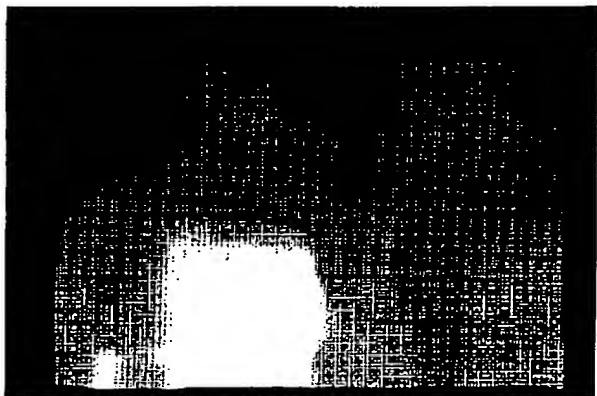
【図 2 2】



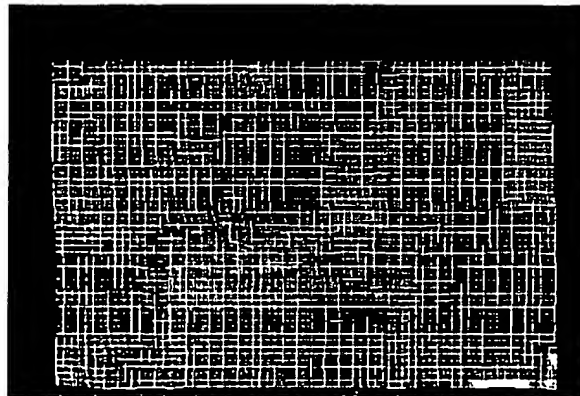
【図 2 3】



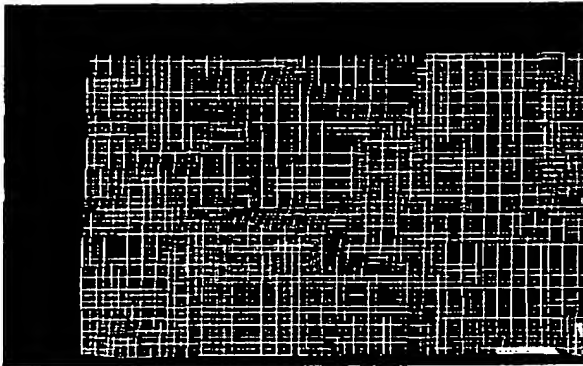
【図 2 4】



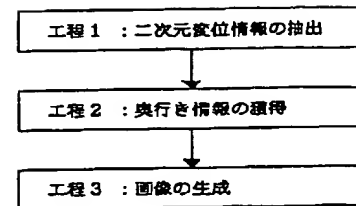
【図 2 5】



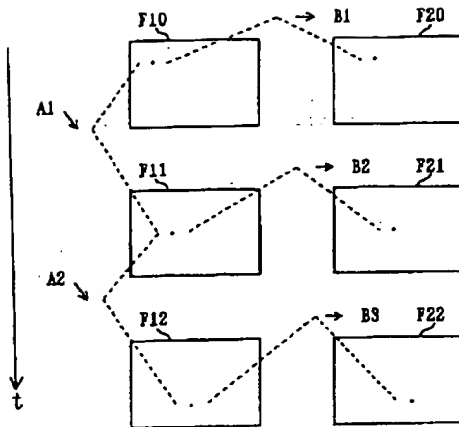
【図26】



【図27】



【図28】



フロントページの続き

(72)発明者 杉本 和英
大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内

(72)発明者 荒川 勉
大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内

(72)発明者 鈴木 治
大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内